The Time Course of Processing External and Internal Facial Features in Face Matching Tasks

Thesis presented to the Faculty of Arts of the
University of Zurich for the degree of
Doctor of Philosophy
by

Bozana Veres-Injac
of Zurich/ ZH

Accepted in the fall semester 2008 on the
recommendation of Prof. Dr. Wolfgang Marx and
Prof. Dr. Heiko Hecht

Zurich, 2008

Contents:

Preface

Faces are extraordinary rich sources of information. From only one glance we can determine important properties of another person: who he or she is, age, sex, gaze direction, and emotional state. This ability is subject to learning experience, and is refined over many years (Diamond & Cary, 1986). The development of expertise in face recognition is accompanied by an increasing relevance of configural information, which holds the specific spatial relationships among facial features (Schwarzer & Zauner, 2003; Mondloch, Grand, & Maurer, 2002; Mondloch**,** Dobson, Parsons, & Maurer, 2004). Moreover, the expert level of face recognition is biased by the long history of perceptual experience with upright faces, and is therefore quite orientation sensitive. We are good in noting fine details and changes in faces when seen in their normal orientation, but when turned upside down, we do not even perceive grotesque deteriorations of the facial whole caused by wrong orientations of mouth, nose or eyes (Bartlett & Searcy, 1993). Stimulus inversion has a stronger impairing effect on face recognition than on common object recognition (for a review see Valentine, 1988). Differences in processing of upright versus inverted faces are taken to reflect differences between special face-expert and generic object processing (Fahra, Tanaka, & Drain, 1995). The very fact that faces are so special to us makes them a fascinating test case for many of the central questions of cognitive psychology, which have been pursued over the last four decades of research.

In the three studies presented here the special contribution of external (hair, face and head outline, ears) and internal (eyes, eyebrows, nose, mouth) facial features to face perception is examined. Focusing on the effects of inversion and viewpoint in the first two studies (Chapter 3 and 4) we were able to postulate two different processing paths involved in handling internal and external feature information during processing of natural face stimuli. The two kinds of features are found to tap separate functional entities of the cognitive system, where internal features are supposed to be processed predominantly by the face recognition system, and external features by the object recognition system. In Study 3 (Chapter 5) the contribution of internal and external features to face perception was further studied with a masking approach apt to discriminate featural and configural modes of processing by their specific timings. Hypothetical processing stages and their temporal order could be revealed, based on evidence for a rapid availability of information provided by external features, whereas information provided by internal features and by integration of both sources are shown to become available at later moments in time. Viewed together, our data give valuable hints at the existence of distinct modes of face processing, and contribute to the current debate about the role and the timing of featural and configural information in face perception.

The thesis is organized as follows. Chapter 1 gives an introduction to the field by reviewing major findings in face perception, and introducing basic research paradigms. In Chapter 2 the main research questions and the experimental rationale are defined. Chapters 3 – 5 contain the three experimental studies that were conducted in order to address processing of internal and external features, and its timing. In Chapter 6 the results of all three studies are discussed together, and major conclusions are drawn from the integration of results. The thesis closes with concluding remarks (Chapter 7), which also gives an outlook to promising directions of future experimental work.

1. Introduction

1.1 Parts and wholes in face perception

Faces belong to the category of homogenous objects, which means that all faces share the same basic parts in the same basic arrangement. Diamond and Carey (1986) distinguish between three kinds of information that can be used in face recognition: isolated features, first-order relations and second-order relations. At the starting point there should be a definition, at least an accepted usage of these terms, since, in the literature, there is no consensus about terminology. Isolated features can be specified without reference to other parts of the stimulus. Here, featural information refers to isolated facial features as in everyday use: hair, eyebrows, eyes, nose, mouth, cheeks and chin. First-order relations refer to the relations among these basics facial features which are shared by all faces, reflecting the typical configuration that defines a face. For example, in faces the horizontally aligned eyes appear above the nose, which is itself above the mouth. Second-order relations refer to the set of spatial relations that characterizes the specific arrangement of the basic features of an individual face. The sensitivity of human observers to fine changes in this type of configurational information makes humans so good in discriminating among individuals, and in detecting changes of emotional states.

A central issue in face perception research is how faces are analyzed and represented in memory. In the nineteenth century, Francis Galton (1879) suggested that the relations among facial features may be more crucial for face recognition than the individual features. By superimposing photographs of many different persons he got a single resultant composite portrait, one that represented no man in particular, but a hypothetical man possessing the average features of the given group. Each of the original faces had a likeness with this prototype portrait, independent of the characteristics of its individual facial features. He suggested that it is this kind of likeness that determines face perception, indicating that facial features are integrated into a holistic percept, and that comparison of individual features is seemingly of minor importance. Although very few attempts have since been made to confirm Galton`s claim experimentally, it is widely accepted that faces are processed holistically.

In modern face perception research there are four main approaches proposing different kinds of mental representations and processing algorithms underlying face perception: 1) the holistic hypothesis, 2) the configurational hypothesis, 3) the featural hypothesis, and 4) the norm hypothesis. The theoretical background as well as some empirical evidence for each hypothesis is described in the following section, which is supplemented by an alternative approach.

1.1.1 The holistic hypothesis

According to this hypothesis we perceive and remember faces as unparsed perceptual wholes (i.e. a facial Gestalt) in which featural and configurational information are not explicitly represented. Representations of whole faces are seemingly more easily and quickly accessed than are representations of parts (Carey & Diamond, 1994). Further, faces appear to be recognized as relatively undifferentiated wholes (Ellis, 1975, Tanaka & Farah, 1993, Fahra et al., 1995). Hence, the central claim of the holistic hypothesis is that the perceptual wholeness of faces is difficult to break down into parts without hampering perception and remembering of facial stimuli.The whole face is assumed to be the “natural” unit of analysis.

Evidence of holistic face perception comes from two basic findings. One is the composite effect, which shows that it takes longer to identify a person who appears as the upper and lower part of a chimeric face when the two parts are aligned, compared to when the halves are misaligned. A chimeric face is represented as a face sectioned along the horizontal midline, whose upper part shows one person and the lower part another person. Identification of the composite face as an unimpaired whole interferes with identification of single parts (Carey & Diamond, 1994; Endo, Masame, & Maruyama, 1989; Hole, 1994; Young, Hellawell, & Hay, 1987). An example of aligned and misaligned chimeric faces is shown in Figure 1.1.

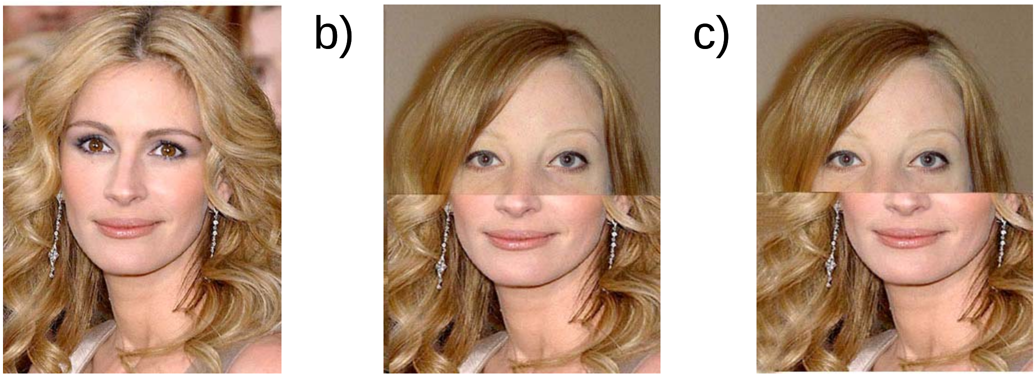

Figure 1.1: The lower part of an original face - here Julia Roberts (a), is easier to recognize as identical to the original in the misaligned chimeric face (c) than in the aligned chimeric face (b). If aligned, two halves of different faces tend to fuse into an integrated new face (b).

The second finding is the better identification of a facial feature (e.g. a nose) when it is seen in the context of a face that has previously been viewed, compared to seeing the part in isolation. This advantage of identifying the-part-in-the-whole is not found with non-facial objects, and seems to be specific of faces (Farah, 1996, Farah et al., 1995, Tanaka & Farah, 1993). According to the holistic approach, face processing occurs in a single step by simultaneous processing of all facial features (Bradshaw & Wallace, 1971) including their spatial relations (Rhodes, 1988; Sergent, 1984; Tanaka & Sengco, 1997).

1.1.2 The configurational hypothesis

The configurational hypothesis proposes that we perceive and remember faces by processing two kinds of information, featural and configural, but with more weight on configural information (Bartlett & Searcy, 1993; Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993; Searcy & Bartlett, 1996).

According to Diamond and Carey (1986) configural information, or second-order information, refers to the specific spatial distances among the basic facial features that define an individual face (see 1.1). This implies that faces also have a prototypical spatial organization of their basic features (first order information) which is not true for all classes of objects. For example, landscapes or random textures do not contain a prototypical arrangement of parts. Therefore, for discriminating among members of these classes of objects the observer must resort to other sources of information. Configural, or second order information is particularly inherent in faces. Moreover, Diamond and Cary (1986) proposed that processing of configural information is the major cause of the face inversion effect (see below), and argued that the inversion effect does not index a perceptual process unique to face perception. Rather, the authors have shown that the same perceptual process underlies identification of highly similar stimuli by viewers that are experts in the particular stimulus domain (e.g. dog recognition by dog experts). So, the use of configural information seems to require knowledge of the general organization of the class of objects under study, and proves to be a powerful means of discrimination among individuals if this knowledge is provided.

Rhodes (1988) defined second-order feature parameters (or configural information) in a similar but resticted way, assuming it as characterizing the position of eyes, the spatial relations among all internal features (but not to external features), and chin shape. For determining facial similarity not only configural information, but also featural information, namely the characteristics of facial features, were found to be important. However, in some later works Rhodes and her colleagues (Rhodes, Brennan, & Carey, 1987; Rhodes & McLean, 1990) shift the focus away from coding spatial relations within a face, and stress the relevance of coding spatial relations between a face and a so called “norm”. This approach is outlined in the section 1.1.4.

1.1.3 The featural hypothesis

The featural hypothesis proposes that faces are perceived and represented as the sum of features (Garner, 1978). Facial features are supposed not to interact with each other, and the overall impression of a face can be understood as the sum of the independent perceptual impressions obtained from each single feature. Smith and Nielsen (1970) have demonstrated that the time needed to discriminate pairs of different schematic faces decreased as the number of differences between the faces increased, pointing to an inclusive decision rule. Further evidence for parallel and independent featural processing was also obtained in other studies (Walker-Smith, 1978; Tversky & Krantz, 1969, Macho & Leder, 1998).

Research on features saliency provided evidence that features may be weighted differently, according to their relevance, but are handled as distinct entities (Ellis, 1975; Shepherd, Davies, & Ellis, 1981; Rakover & Teucher, 1997). Hair, face outline, eyes and mouth (not necessarily in this order) have been shown to determine perception and memory of faces (McKelvie, 1976; Shepherd et al., 1981; Fraser, Craig, & Parker, 1990). The nose plays a minor role in perceiving and remembering faces in frontal view, but a particular and highly distinctive nose shape can be more important than eyes or mouth for recognizing faces in profile view (Roberts & Bruce, 1988).

A major line of proof for the featural hypothesis comes from experiments which aim at showing that different facial stimuli become perceptually available at different moments in time, indicating that face processing can be viewed as a sequence of feature related events rather than capturing a stimulus configuration at one moment in time (Flavell & Draguns, 1957; Bachmann, 1991; Carbon & Leder, 2005). Indeed, some studies have demonstrated the relevance of piecemeal feature processing in face recognition by showing that the speed of recognition differs significantly among different types of facial features (Bartlett, Searcy, & Abdi, 2003; Rakover & Teucher, 1997).

1.1.4 The norm hypothesis

According to the norm, or prototype hypothesis, featural and configural facial information is represented in a cognitive system as the deviation from a prototype in the form of an abstract distance measure. The facial norm or prototype is a representation within a (human or machine) cognitive system built by learning via extensive exposure to, or interaction with, a large number of faces. Encoding a face is conceptualized as a comparison to a norm, with the representation being defined in terms of its deviations from that norm along fixed and predefined dimensions. These deviations emphasize the featural or the configural quality of a given face with the respect to the norm (Rhodes et al, 1993; Rhodes et al, 1987; Rhodes, Carey, Byatt, & Proffitt, 1998; Valentine, 1991).

There are three main lines of evidence supporting the norm hypothesis. The first one is the effect of general type accordance, or distinctiveness on face perception. Bartlett, Hurry, and Thorley (1984) have showed that subjects can reliably judge whether a face is a more or less typical face, or a particular instance. Moreover, in forced choice comparison with jumbled non-faces, typical faces are recognized to be faces more quickly than atypical faces (Valentine & Bruce, 1986a; Valentine, 1991). Also the attractiveness of a face depends on the degree to which its configuration matches to an average faces (Langlois & Roggman, 1990; Langlois, Roggman, & Musselman, 1994; Rhodes & Tremewan, 1996).

A second line of evidence suggests that new faces may be encoded in terms of features that deviate from the norm. In recognition memory tasks with unfamiliar faces, atypical faces and faces with distinctive features are remembered better than typical faces (Bartlett et al, 1984; Light, Kayra-Stuart, & Hollander, 1979; Winograd, 1981; Johnston & Ellis, 1995a; Leder & Burce, 1998). To establish the distinctiveness of faces Brennan (1985) measured the locations of a large number of meaningful points in facial images (for example left corner of upper lip or the starting point of right eyebrow). Bruce, Burton, and Dench (1994) revealed that the rated distinctiveness of faces can be accounted in terms of their physical deviation from the norm.

A third line of evidence comes from studies with caricatures. In caricatures facial features appear exaggerated with respect to the norm that was used. Caricatures are found to be recognized as good as or better than the corresponding original faces, and better than anti-caricatures, in which differences between face instance and norm is reduced (Benson & Perret, 1991, 1994; Rhodes & Tremewan, 1994; Mauro & Kubovy, 1992; Stevenage, 1995).

1.1.5 Critical spatial frequency bands

Two types of information have been shown to be more important for the perception and identification of faces than for other complex objects: configural relations and critical bands of spatial frequencies. Harmon (1973) and Harmon and Julesz (1973) suggested that the main facial information required to identify a face is located in the range of low spatial frequencies. These frequencies do not show the details of a face, but its general outline and configuration. The large amount of research that followed the studies of Harmon and Julesz did not support this hypothesis. However, it is generally accepted that recognition of faces is preferentially supported by limited ranges of spatial frequencies. Although different experimental paradigms led to inconsistent results, there is now general agreement that this critical band lies between 6 and 12 cycles per faces-width (Bachmann, 1991; Costen, Parker, & Craw, 1996; Fiorentini, Maffei, & Sandini, 1983; Näsänen, 1999). This band of spatial frequencies is said to be critical because face images that have been filtered such to contain only information within this band are better recognized than filtered faces containing spatial frequencies either below or above this band. In contrast, object recognition is generally less affected by spatial frequency filtering, i.e. cultural objects can be recognized on multiple spatial scales (Biederman & Kalocsai, 1997).

There is a number of plausible possibilities why such a critical band should exist. The advantage of medium frequencies may reflect the selective use of spatial frequency information by a specialized face recognition system. Another specialized system, for example the one dedicated to the recognition of letters, may use spatial frequency information of other ranges (Solomon & Pelli, 1994). The other possibility is that the medium range is critical for encoding configural information, which is essential for face recognition. However, many studies showed that, in the medium frequency band condition, the ability to extract and discriminate not only configural but also featural information was superior, leading to the conclusion that spatial frequencies critical for face recognition do not preferentially contribute to configural encoding (Wenger & Townsend, 2000; Nagayama, Yoshida, & Toshima, 1995; Boutet, Collin, & Faubert, 2003). Viewed together, the results suggest that lower spatial frequencies (<8 cycles per faces) clearly support extraction of configural cues, whereas high spatial frequencies (>32 cycles per faces) support extraction of local features in face recognition (Goffaux & Rossion, 2006; Goffaux, Hault, Michel, Vuong, & Rossion, 2005; Collishaw & Hole, 2000).

1.2 Face inversion effect (FIE)

Very early in the history of face recognition research there was found clear evidence that inverted faces are more difficult to recognize than upright faces (e.g. Goldstein, 1965, Hochberg & Galper, 1967). As reflected by the vast amount of studies dedicated to the face inversion effect (FIE) since then, this effect really has become an intriguing topic. There is a strong claim that inversion affects face processing *disproportionately* compared to processing of other objects which are also usually only seen upright (Yin, 1969). Because of the strong impact of Yin's study on successive research and its relevance for the quarrel about featural and configural modes of processing this study is briefly outlined.

Yin (1969) compared recognition memory for photographs of faces with other stimuli, which are also predominantly seen upright, but rarely in other orientations: houses, airplanes and stick figures of men in motion. A group of stimuli was initially presented upright or inverted. Recognition memory was tested by a two-alternative forced-choice procedure, using items in the same orientation as the initially presented objects. When the stimuli were presented and tested upright, faces were better recognized than other stimuli. However, if presented inverted, face recognition deteriorated dramatically. This was the first study showing that faces compared to other mono-oriented objects are more affected by inversion. Yin interpreted his findings as evidence for face-specific processing distinguishing two aspects: first, experience with an object in its usual orientation, which is not specific just for faces, but concerns all mono-oriented objects, and, second, a factor specific for faces. The results of the study were replicated many times, with different stimuli and across a wide variety of experimental conditions (Scapinello & Yarmey, 1970; Yarmey, 1971; Toyama, 1975, Carey & Diamond, 1977; Diamond & Carey, 1986). However, the question of whether face processing is “special” with respect to inversion has remained a central theoretical issue in the literature up to now.

1.2.1 Are faces special?

In the seventies Ellis (1975) analyzed three issue of evidence used to support the view that face processing is special: ontogeny of face recognition, the clinical evidence with prosopagnosia patients (the inability to recognize familiar faces) and the effect of inversion. He found that the analysis of ontogeny does not provide unique support for the view that young infants’ interests for faces are innate rather than acquired. Similarly, analyzing the other two evidence line he did not find “any unambiguous evidence that faces are handled by a special and specific recognition system” (p.424).

Hay and Young (1982) proposed that the question about face specialty can be split into two different questions. First, is there a part of the brain which is specific to face recognition, involved only in face processing? Second, does face recognition involve a unique process? The authors proposed that only uniqueness of processes would imply that face recognition is qualitatively different from recognition of all other visual categories of objects. The position they support assumes a specific mechanism which does not employ a unique process (Hay & Young, 1982). Accordingly, they interpret evidence from prosopagnosia as supporting specificity rather than uniqueness of face recognition.

Further evidence against specialty in face processing came from Diamond and Carey (1986). They studied the effect of inversion upon recognition of faces and dogs in naïve subjects and dog experts. They explored whether faces are uniquely represented in memory in terms of distinctive features which are particularly sensitive to inversion. The univocal answer of Diamond and Carey is that faces are *not* special stimuli. In their study the authors showed that perceiving other classes of visual stimuli (dogs) is as sensitive to orientation as face perception, provided there is some expertise with this class of stimuli. Diamond and Carey (1986) conclude from their data that a large inversion effect will emerge whenever three conditions hold. First, the members of the stimulus class have to share the same configuration. Second, the members of the class can be individuated on the basis of second-order relational features (configuration among features). Third, subjects have to be experts in using such features to distinguish between the members of the same class. In other words, there are strong differences among the perceptual strategies used by novices and experts in face perception. Experts are supposed to have the ability to distinguish face instances on the basis of subtle differences in configural information, whereas novices prefer to rely on isolated features. The expert's use of configural information is assumed to be hampered by inversion, so that in the inverted condition both experts and novices rely upon isolated features, while in the upright condition experts exploit configural information. This account predicts better performance of experts in recognizing upright photographs of e.g. dogs, as used in the experiment, and same performance for experts and novices with inverted stimuli. However, Diamond and Carey (1986) did not find any significant effect of expertise on recognition of upright dog photos, and novices were found to perform as well as experts. Contrary, novices performed better than experts at recognizing inverted dog pictures. Although the authors attribute these problems to a failure to control the age of subjects (experts mean age = 64 compared to undergraduate novices), there is little support for the assumption that experts change their perceptual strategy with inverted stimuli, but seemingly used configural information in both conditions. Similar findings have not yet been published, and recent efforts to replication failed (Robbins & McKone, 2007). Another recent study using fingerprint stimuli with normal observers and experts also failed to find a significant inversion effect mediated by object expertise (Busey & Vanderkolk, 2005).

A study conducted by Valentine and Bruce (1986b) also yields evidence against the “expert” hypothesis in the domain of face recognition. They examined the effect of inversion upon recognition with classes of faces differing in the degree of familiarity: own and other race faces. According to the expertise hypothesis greater familiarity with a class of objects enhances the inversion effect. Therefore, recognition of own race faces is expected to show a larger inversion effect than recognition of other race faces observers are not familiar with. However, the opposite result was obtained: the inversion effect was larger for unfamiliar, other race faces than for the faces of the own race. This data suggest that Yin`s result cannot be explained in terms of stimulus class familiarity and directly contradict the expertise hypothesis proposed by Diamond and Carey (1986).

1.2.1.1 Neuropsychological evidence

Early evidence that there may be specialized neural regions for face perception came from cases of prosopagnosia, the selective loss of face recognition abilities in patients with a focal brain damage (e.g. Bodamer, 1947; Meadows, 1974). However, the first neuropsychological evidence of such specialization came from the discovery of face-selective cells in the temporal cortex of macaques (Gross, Roche-Miranda, & Bender, 1972). Since this pioneering work many studies have demonstrated that “face cells” may be tuned to certain facial attributes such as personal identity (Yamane, Kaji, & Kawano, 1988), expression (Hasselmo, Rolls, & Baylis, 1989), viewpoint (Perret et al, 1991), or part of a faces (Perret, Rolls, & Caan, 1982, Yamane et al, 1988).

In the early 1990s, PET (position emission tomography) studies demonstrated activation of the ventral visual pathway, especially the fusiform gyrus in a variety of face perception tasks (Haxby et al, 1999, Sergent, Ohta, & MacDonald, 1992). fMRI (functional magnetic resonance imaging) studies of the specificity of these cortical regions for faces *per se* began in the mid-1990s, showing that fusiform regions respond more strongly to faces than to letter strings and textures (Puce, Allison, Asgari, Gore, McCarthy, 1996), flowers (McCarthy, Puce, Gore, & Allison, 1997), and other stimuli including mixed everyday objects, houses and hands (Kanwisher, McDermott, & Chun, 1997). The most consistent and robust face-selective activation was located on the lateral side of the min-fusiform gyrus in the region named the “fusiform face area” or FFA (Kanwisher et al., 1997).

Although the “face specialization hypothesis” has remained controversial, there is substantial evidence that face perception may be implemented in its own specialized cortical network that is not shared with many, if any, other cognitive functions (for a review see Kanwisher & Yovel, 2006). Contrary, according to the expertise hypothesis face-specific mechanisms are specialized not for processing faces *per se*, but rather for distinguishing between exemplars of a category that share the same basic configuration, and for exemplars for which the subject has acquired substantial expertise. This means that FFA may respond not only to faces, but also to other kinds of objects which contain configural information.

The strongest neuropsychological evidence for this hypothesis come from the study of Gauthier and colleges (Gauthier, Williams, Tarr, & Tanaka, 1998) on the basis of fMRI studies in which subjects undergo extensive training in the laboratory on novel perceptual objects called “Greebels”. They scanned subjects looking at faces and Greebles, and report that the difference between upright and inverted Greebles in the FFA region increased through Greeble training. However, there are several problems precluding this data from being interpreted as clear evidence for an expertise effect in FFA. First, the authors do not report data about the magnitude of response to upright Greebles and upright faces after training, which should serve as a control to the temporal change. Second, Greebles are not appropriate artificial stimuli, since they resemble faces and/or bodies to large amounts. Third, the FFA in this study was defined very large, and therefore it can not be excluded that any training effects on Greebles may arise from the other neighboring regions (e.g. body selective fusiform body area).

In recent studies avoiding these experimental flaws (Moore, Cohen, & Ranganath, 2006; Yue, Tjan, Biederman, 2006, Op de Beek, Baker, DiCarlo, & Kanwisher, 2006) subjects were trained for many hours on fine-grained discrimination between exemplars of novel objects that do not resemble faces or bodies. None of these studies found a significant increase in the response of FFA for trained compared with untrained objects, but all found significant training-inducted increase of responses in a nearby region called the lateral occipital complex (LOC), which is responsive to the object shape in general, but not for faces in particular. Thus, expertise effects are not restricted to the FFA, but rather outside of FFA (Rhodes, Byatt, Michie, & Puce, 2004), and are not face specific.

Moreover, there is neuropsychological evidence that inverted faces do not engage face perception mechanisms, but are processed by mechanisms dedicated to the perception of other objects. In patients with prosopagnosia, which is a selective impairment of face recognition, the recognition of inverted faces can be relatively normal, suggesting that inverted face perception may be mediated by intact object perception mechanisms. In fact, some prosopagnosic patients perform worse on tasks with upright than inverted faces (Farah, Wilson, Drain, & Tanaka, 1995), suggesting that upright faces evoke erroneous processing by a damaged face perception system. Inverted faces, on the other hand, do not evoke processing by the damaged face recognition system, thereby allowing the intact object recognition system to operate without interference. Additional evidence that the face perception system cannot process inverted faces effectively comes from the studies with object agnosia patients. Perception of upright faces was normal, but recognition of inverted faces was seriously impaired (Moscovitch, Winocur, & Behrmann, 1997). Studies using fMRI revealed that the effect of face inversion was an increased response in ventral extrastriate regions that respond preferentially to other classes of specific objects (e.g. houses). In contrast, inversion of these objects did not produce a similar effect in face-selective regions (Haxby et al., 1999).

Taken together, the neuropsychological and behavioral data do not seem to provide convincing evidence for the expertise hypothesis and indicate that mechanisms of face-specific processing and object related processing can be localized by combining imaging techniques with proper experimental paradigms.

1.2.2 Configural and featural information in inverted faces

As discussed above (see 1.1) there is general agreement about two sources of information involved in face processing: featural, relying on piecemeal features (e.g eyes, nose, mouth), and configural, defined as spatial relations among these features (i.e. distances between eyes, or eyes and mouth, etc.). The dual-mode hypothesis postulates that two different visual processes are required by upright and inverted faces (Bartlett & Searcy, 1993). It is maintained that configural information processing is much less efficient in inverted faces, possibly even entirely disabled. The perception of inverted faces must therefore proceed by a laborious, serial-part-by-part analysis of the individual facial features. The dual-mode hypothesis implies that configural information may be less important input element for the processing of inverted than upright faces. Indeed, there is empirical evidence that inversion impairs perception of configural relations in faces, but leaves featural analysis unaffected.

Young et al. (1987) combined top and bottom halves of different faces, creating so-called chimetric faces (an example is shown in the Figure 1.1). They have found that fusing the two halves of different famous faces slowed reaction times for the recognition of the identity of the halves, compared to conditions were the halves were horizontally offset (misaligned). This effect was not seen in inverted faces, indicating that configural fusion does not occur in inverted faces. Kemp, McManus, and Pigott (1990) showed that small shifts of the eyes upwards or nasally were harder to discriminate in inverted than upright faces. Searcy and Bartlett (1996) made faces grotesque by either changing featural aspects (blurring the pupils, blackening teeth) or by distorting the facial configuration. When presented inverted, faces which were made grotesque through configural distortions appeared much more similar to the unaltered version while with featural distortions faces still appeared grotesque. Similarly, Leder and Bruce (1998) made faces more distinctive in appearance either through changes in featural aspects (such as bushier eyebrows, darker lips, etc.), or through changes in configuration (such as narrower interocular distances, shorter mouth-to-nose distances, etc). Distinctiveness effect produced by configural properties vanished when faces were presented inverted compared with faces either with featural alternation or the same faces in upright orientation. In a recent approach Leder and Bruce (2000) tested directly whether critical configural information (interocular distances or nose-to-mouth distances) is necessary for the inversion effect in face recognition to occur. They used schematic faces, which differed in terms of configural properties from each other, but consisted of identical features (eyes, nose, mouth, etc.). Participants learned to name each face in a familiarization phase. A series of different recognition experiments revealed that the configural information was necessary for the occurrence of the inversion effect. Schwaninger and Mast (2005) tested their participants in an individual discrimination task on faces presented at multiple orientations (0, 30, 60, 90, 120, 150, 180 degrees) while manipulating distances between facial features (configural information). They found a massive drop of performance and RT increase between 60^{o} and 90^{o} when distances between features were manipulated (Schwaninger & Mast, 2005). The data are congruent with a recent study by Rossion and Boremanse (2008) showing that holistic or configural face processing is particularly impaired by rotations of about 90^o. The results of all these studies suggest that configural information plays a stronger role in the processing of upright than of inverted faces.

If the face inversion effect is caused by the destruction of configural information, limiting the observer to the use of featural information only, then the presentation of isolated inverted features should not result in a reduction of recognition, since there is no configural information to be extracted from an isolated facial feature either in upright or in inverted orientation. Rhodes et al. (1993) tested this hypothesis by presenting eyes or mouth as isolated features, without being displayed in a facial context. Their results revealed that the inversion effect was caused by spatial relations (spacing between features), but also indicate a large inversion effect for features (eyes, mouth) when presented in the context of face outlines. Further, the inversion effect disappeared when the features were presented free of facial context. The notion that recognition of facial features is unaffected by inversion was challenged by Rakover and Teucher (1997), who found that recognition of the forehead, eyes, nose, mouth and chin is less accurate in inverted than in upright orientation. They suggest that configural information is not necessary for obtaining an inversion effect with a whole face, and that in the recognition of such a whole face, configural information extracted from an upright face is less important than featural information. This hypothesis was corroborated by Moscovitch and Moscovitch (2000) who found that inverted faces are not processed by piecemeal analysis of their features, since the spatial relations among features, as well as their orientation are necessary for face recognition. Considering facial features, Nachson and Shechory (2002) made a distinction between external (e.g. hair, ears, facial outline) and internal (e.g. eyes, eyebrows, nose, mouth) features. In two tasks (pair and multiple-choice matching) they studied whether or not inversion affects recognition of full faces, as well as isolated internal and external features. In line with previous studies, the data consistently showed that in both tasks inversion similarly impaired recognition of both full faces and facial features. Barton, Keenan, and Bass (2001) have found an inversion effect for the perception of feature position, as well as for feature colour, which was highly modulated by regional salience. In the less salient bottom region of faces the inversion effect was stronger (e.g. inversion effect for mouth position compared with eye position), and was affected by viewing duration, where longer time diminished the FIE. Taken together, these younger studies suggest that the inversion effect does not result from a shift of configurational (holistical) towards feature-by-feature perception. Rather, they suggest that processing of both configural and featural information is orientation sensitive, the latter at least to some degree.

1.3 Viewpoint dependency or independency?

Moving around in the world, we encounter a large number of objects and faces, which we recognize immediately and effortlessly. This is remarkable by itself, but it becomes even more intriguing when one takes into account the myriad of views that can be generated from each single object or face. The ability to recognize objects on the basis of its variable image instances (visual object constancy) has evolved into one of the most intensely debated topics in vision science (Lawson, 1999). The issue has been investigated with a great variety of research methods. The results on viewpoint dependency are, however, rather variable and inconclusive. In the following section we discuss two major classes of theoretical accounts, relying on the different assumptions about the nature of the internal representation and the relative importance of the observer`s viewpoint (for a more detailed review see Tarr, 1995; Vecera, 1998).

1.3.1 Matching of objects from different viewpoints

There are two main classes of theories accounting for view effects in the perception of objects: theories which claim that object representation are somewhat “concrete” and, therefore, object recognition must depend on viewpoint, and structural description theories which claim that object representations comprise descriptions of the relations among object parts, which are viewpoint independent.

1.3.1.1 Viewpoint dependency

Theories of viewpoint dependency propose that there is a “canonical view”, in which objects are quickly and accurately named and recognized (Palmer, Rosch, & Chase, 1981). Increasing the angle of rotation relative to the canonical view monotonically increases reaction time, an effect often interpreted as evidence for mental rotation, assumed to compensate the discrepancy between the canonical view and a new object view (Shepard & Metzler, 1971; Palmer, 1983; Shepherd & Cooper, 1982; Tarr & Pinker, 1989). As an alternative to mental rotation, neural extrapolation between more than a single canonical view was introduced by supporters of neural networks approaches to object perception (Poggio & Edelman, 1990; Riesenhuber & Poggio, 2000). There is a large body of evidence for viewpoint dependent recognition, in that performance is a function of misorientation relative to the learned view(s), with a variety of conditions and types of stimuli (e.g. Edelman & Bülthoff, 1992; Farah, Rochlin, & Klein, 1994; Rock & Di Vita, 1987; Tarr & Pinker, 1989, 1990).

A well-established method to elicit a viewpoint-dependent process is asking subjects to discriminate between an object and its mirrored counterpart. Shepard and Metzler (1971) demonstrated that, in this case, response times constitute a linear function of the angular difference (AD) between the two presented objects, indicating that subject mentally rotated one of the objects, and that rotation about a angular unit amounts to a fixed time for rotation. Since Shephard’s and Metzler’s experiment the increasing response time curve in such "handedness" tasks has been replicated frequently (for reviews see Kosslyn, 1994). There seems to be a consensus that a handedness task elicits the mental-rotation process and it remains perhaps the clearest example of viewpoint-dependency in visual object perception.

Bülthoff and Edelman (1992) trained subjects to recognize different views of unfamiliar "paper-clip-like" objects. In a later testing stage it was shown that performance was best for the learned views, but also that different new views elicited different patterns of generalization. Specifically, generalization performance was better for new views interpolating trained views, than for new views extrapolating the trained views, and worst for new views orthogonal to the trained views.

One theoretical interpretation of viewpoint dependence is that objects are stored in memory as collections of discrete views (Tarr & Pinker, 1989, Bülthoff & Edelman, 1992). Different mechanisms have been proposed to allow new object views to be generated from stored views – alignment (Ullman, 1989), linear combinations of views (Ullman & Basri, 1991), or interpolation between stored views (Poggio & Edelman, 1990). Linear combination and interpolation among stored views are mechanisms that can be assumed to operate faster than mental rotation.

1.3.1.2 Viewpoint independency

Theories of viewpoint dependency contrast with structural, model-based accounts which propose that objects are represented as sets of parts and their relations (e.g. Biederman, 1987; Marr & Nishihara, 1978). In all theories maintaining viewpoint independency the extraction of invariant features plays an essential role in constructing an object-centered representation. A strong assumption of structural theories implies viewpoint independence over a wide range of viewpoints, because the component parts of objects are assumed to be stored in terms of "non- accidental" properties. In structural theories objects are described in terms of the spatial relations among their parts, and the parts are described of terms of parallelism, curvature (concavity/convexity), connectedness, and termination. The latter properties are “non-accidental” in the sense that they describe object parts independent of the way we look at them. As long as we look at an object from a normal, and not from an accidental viewpoint (e.g. looking at a tube directly from the front would hide its 3D structure, therefore this accidental view would preclude to extract length and curvature of the tube, and we could erroneously confuse it with a ring) we are able to extract the non-accidental properties of its components. Non-accidental properties allow to recover 3D shapes from their 2D views, as long as the same parts remain visible over the range of considered views (Lowe, 1987). By introducing a diagnostic part which remains invariant across rotations in depth it is possible to make explicit the process of employing invariants in recognition (for a review see Vecera, 1998). As there is no need to presume some sort of transformation, these theories predict that there should be no substantial effect of viewpoint manipulation on processing times.

The use of invariant features is apparent in the behavioral response pattern: both Eley (1982) and Biederman and Gerhardstein (1993) found no effect of changes in viewpoint. In their priming study which required the speeded naming of line drawings of depth-rotated views of familiar objects they reported that the prime had little effect on target naming. They claimed that in everyday situations, object recognition is largely invariant to depth rotation, and suggested that the strong effect of depth rotation reported in several studies (Edelman & Bülthoff, 1992; Rock & Di Vita, 1987; Tarr & Pinker, 1989) was due to testing novel, or artifical objects.

Biederman and Gerhardstein (1993) summarized three conditions for viewpoint independence in part-based theories of object recognition: 1) an object must be decomposable into its parts, 2) different objects must be distinguishable by different parts, and 3) the same part- based description must be recoverable from different viewpoints.

It is important to note that there is usually more information available from a particular image (shape, part, color, texture, or other derived cues) that can be diagnostic for a particular task or categorization of the object. Viewpoint dependency can emerge when diagnostic information is available in different views. The impact of such “accidental surface properties” is often underestimated in structural theories. For example, the structural descriptions of a rose and a tulip are quite similar, but both flowers are easily distinguished with the possible colors they can take.

1.3.1.3 Multiple routes to object recognition

Both classes of theories, view-dependent as well as view-independent theories, have provided substantial empirical evidence to their favor. However, when reviewing the relevant literature one instantly notices the wide range of both experimental paradigms and their specific binding to stimuli serving to produce the contradicting data (Hayward & Tarr, 1997). Moreover, results also seem to be correlated with the required level of recognition (Bülthoff, Edelman, & Tarr, 1995; Tarr, 1995). These authors have shown that if the object has to be recognized at the entry level, behavioral measures are relatively unaffected by changes in viewpoint. However, in the case of recognition on subordinate level requiring more subtle discriminations, both response times and accuracy are far more sensitive to viewpoint. Furthermore, differences in the task a subject has to perform (Lawson, 1999) and the specific paradigm that is used (Verfaillie, 1992) can influence which level of representation is tapped. A nice criticism and an integrative view is found in Hayward (2003).

Given this growing amount of conflicting results, it becomes increasingly difficult to hold the assumption of a single, unitary system using only viewpoint or only structural information. For this reason notions such as "multiple routes to object recognition" are adopted more and more frequently in vision science in recent years (Humphreys & Riddoch, 1984; Lawson, 1999). More specifically, there seems to be growing consensus that there is a continuum from viewpoint-dependency to viewpoint-independency, which is in part influenced by stimulus discriminability and the task (e.g. Biederman & Bar, 1999; Edelman, 1995; Hayward & Tarr, 1997; Hayward & Williams, 2000; Van Lier & Wagemans, 1999; Vanrie, Willems, Wagemans, 2001). In short, the key question is no longer *if* object recognition is viewpoint-dependent or independent, but rather *when*, i.e. under which circumstances.

1.3.2 Matching of faces from different viewpoints

To recognize a face from a novel view humans must be able to encode something unique about the face that distinguishes it from all other possible faces, and, they must be able to access this unique information from the novel view. Thus the representation of any particular face in memory must encode the uniqueness of that face, but also be versatile enough to generalize across possible changes in the image characteristics of that face (e.g. viewpoint, expression or illumination). Studying how we accomplish this task is difficult due to the complexity of the visual information observers experience in viewing faces from different viewpoints, and due to multiple ways in that such information can be encoded and represented.

1.3.2.1 A face-space model of face representation

Some models of object and face recognition appeal to the notion of inter-item similarity as a predictor of recognition performance across views. Generally, items that are less similar should be recognized more easily across changes in viewpoint than items that are highly similar (Edelman, 1995, Valentine, 1991, Newell, 1998). When the task involves memory of faces, distinctive faces are recognized faster and more accurately than typical or average faces. Distinctive faces are also less likely to be confused with other faces than are typical faces (Cohen & Carr, 1975; Going & Read, 1974; Light et al., 1979; Valentine & Bruce, 1986a, 1986c). The effect of distinctiveness on face processing can be interpreted by thinking of faces as located in face space. The center of the space is assumed to represent the average value of the population on each dimension. The dimensions of the space will be those that serve to discriminate between faces.

The "face-space" model of face representation in visual memory accounts for the findings on the recognition of distinctive and typical faces (Johnston & Ellis, 1995b; Valentine, 1991). Particularly, this model makes predictions about the recognition of faces based on their dicriminability: the face representations in memory can be considered in terms of locations in a multidimensional feature space. The position of a face representation in face space reflects inter-item similarity so that similar faces will be located in closer proximity than less similar faces. Here similarity is often defined by the physical characteristics of the face which provide a distance measure in an assumed latent space which describes how far an individual face instance is away from the mean of the set of faces, as is done in multidimensional scaling (Benson & Perrett, 1994; Bruce et al., 1994). The underlying principle of the face space model is that recognition performance is determined by the level of inter-item similarity between the faces, which in turn depends on the relative positions of the face instances in a latent facial feature space.

Newell, Chiroro, and Valentine (1999) proposed two ways in which an item-similarity model of face representation can incorporate recognition from different viewpoints. They have suggested that representations in the face space may either be *view-based* or *individual-based*. If the face space was organized by viewpoint, then different views of faces would inhabit different regions of the face space. In this way, the face space is organized according to the physical similarity between the images of each face (Benson & Perrett, 1994, Bruce et al, 1994). The location of an individual`s face relative to the other faces will remain invariant across the sub-spaces, for example a distinctive face will be distinctive from all viewpoints. The separate view-specific representation of the same individual must be bound together. The authors refer to this hypothesis as the view-based account. This is illustrated in Figure 1.2.

Second, the individual-based account proposes that all views of faces may be located in the same face space, with the different views of the same individuals located adjacent to each other. Accordingly, each individual is represented by a clustering of views located together within the same face space. The problem here is that a face space needs to encode the different views of an individual independent of their similarity to other face views.

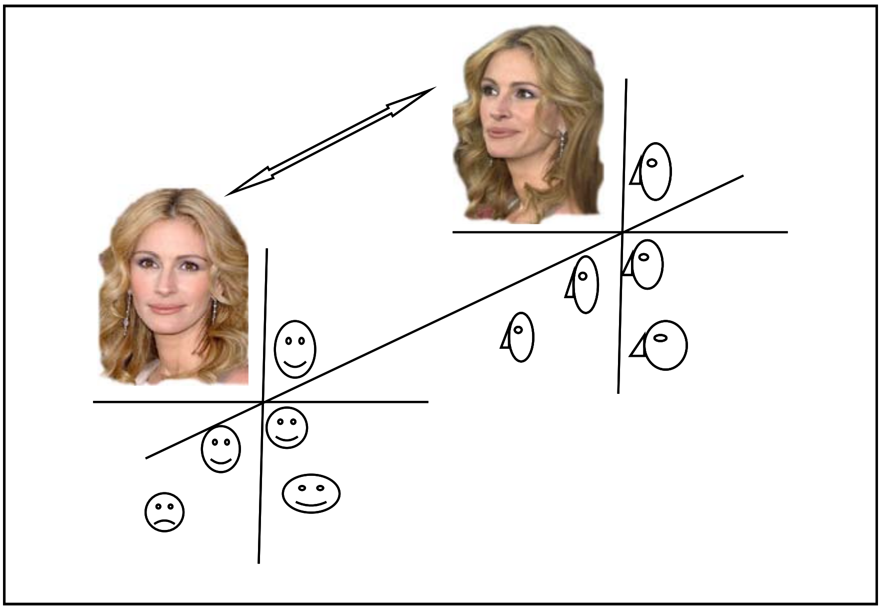

Figure 1.2: A schematic illustration of the predictions made by the view-based account of the face space model (Newell et al., 1999). The different views of a face are located within different view-specific sub-regions of the face space. The distance between two faces within any sub-region is defined by similarities between those two faces. Distinctiveness is retained across the view dimension. The separate view-specific representation of the same individual are associated as indicated by the arrow between two faces.

The authors proposed that this integration process might be achieved as a result of temporal processing at the input stage (Wallis & Bülthoff, 1997, Wallis & Rolls, 1997). Accordingly, images that are presented in close temporal proximity are often associated as belonging to the same object or face, even when small differences in the spatial characteristics of the images occur. Indeed, there is evidence that items presented in close temporal sequence are often associated with each other in visual memory (Miyashita, 1988; Miyashita & Chang, 1988). The authors refer to this hypothesis as the individual-based account. This is illustrated in Figure 1.3. The view-based and individual-based accounts make different predictions on recognition performance across views. If the face space is organized as suggested by the view-based hypothesis, than the recognition performance to novel views should be the same for both typical and distinctive faces.

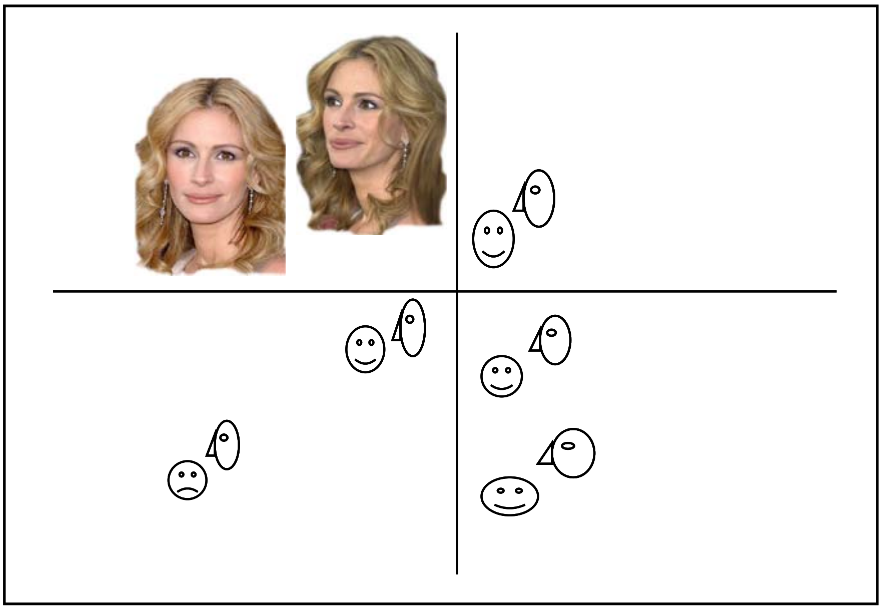

Figure 1.3: A schematic illustration of the predictions made by the individual-based account of the face space model (Newell et al., 1999). Here, the different views of a face are clustered together and are encoded within the same region. The distance between individual clusters reflects the similarity between the individual faces.

Recognition performance is expected to decrease with the increasing difference in viewpoint, but this decrease should be the same for both types of faces because a change of viewpoint would amount to the same physical change on both typical and distinctive faces. In contrast, if the face space is organized as suggested by the individual-based account, then generalization to new views should be defined by the discriminability of that individual`s face. Specifically, novel views of a distinctive face should be recognized more easily than novel views of a typical face. This is expected due to the general organization of the face space: distinctive faces, with all their different view representations, are far apart from their neighboring faces, and, therefore, it would be more likely that novel views of a distinctive face would be correctly associated to the target face. The recognition of a novel view of a typical face is expected to be poor because a new view would be encoded in a densely populated region of the face space.

To test whether the discriminability of a face interacts with the recognition of that face across changes in viewpoint the authors conducted three experiments (Newell et al, 1999). In summary, the results showed that changing the view of a face that is seen before makes recognition more difficult and slower. When all three views of a previously unfamiliar faces have been seen, recognition of ¾ and full-face views was faster and more accurate than the recognition of profile views. Most importantly, a distinctive face is easier to recognize than a typical face, but two views of a distinctive face were not easier to match than were two views of a typical face. This is clear evidence that there is no interaction between the effect of distinctiveness and generalization to a novel view on recognition memory for faces. Instead, the results seem to support the view-based hypothesis assuming that the face space relies on view-centered encodings.

1.3.2.2 ¾ view effect in face recognition

It is well established that certain views of objects are easier to recognize than others. These views are often referred to as “canonical views”. For most objects, a ¾ view is preferred as a canonical view. Palmer et al. (1981) have shown that canonical views have numerous advantages compared to the other views, e.g. they are frequently rated by subjects as more representative than others, are more often created in mental imaginary, and lead to a faster naming responses.

It has been speculated in the face perception literature whether there might be a similar kind of canonical view in which faces are more readily identified and remembered (Bruce, Valentine, & Baddeley, 1987). In a study by Blanz, Tarr, and Bülthoff (1999) it was found that when asked to choose the best view to represent a 3D face in a brochure, most participants chose the ¾ view. As supporting evidence for the ¾ view preference Baddeley and Woodhead (1983) observed that the traditional portraits contained in the National Portrait Gallery in London have been most frequently drawn in ¾ view since the Renaissance period. Although this observation may have overlooked other factors behind this phenomenon, the suggestion that the ¾ view may be a more informative way to convey the characteristics of a face is appealing. A portrait in ¾ view does appear to show more aspects of a face than in other views.

As for objects, the ¾ view is an obvious candidate for a canonical face view, located around the middle point between a full face (or frontal) view and the profile view. Namely, Poggio, and Vetter (1992) have shown that the recognition of a bilaterally symmetric object from a novel view could be achieved if only one nonsingular view of the object is known. If perception “assumes” symmetry, this could generate a symmetric “virtual” view from the only known view, or exploit equivalent information. A face is approximately bilaterally symmetric. ¾ view is a non-singular view from which a symmetric view can be generated. The full-face view, however, is singular.

Single-cell recordings in the macaque superior temporal sulcus (STS) have identified cells which are preferentially tuned to respond to specific views of a head (Harries & Perrett, 1991; Perret, Mistlin, & Chitty, 1989; Perrett et al, 1991). Most of the cells were viewer-centered responding unimodally to one view (either the frontal, the two profiles, or the back views), whereas few cells were tuned to other views of the 360 degree range. The preference for the ¾ view is naturally interpreted in light of Perrett`s findings as the view which elicits the highest total activity from both the profile and full-face neurons. This activation is higher than the response of the individual cells to their preferred view.

In the literature, the ¾ view advantage was often defined in terms of measurable behavioral effects. Specifically, the two types of ¾ view effects can be identified in the literature. The first type of advantage shows that faces learned or tested in ¾ view may generalize better to other views and can be referred as a “”different view advantage”. The second one shows that faces in ¾ view are better identified than other views when the training and test view are identical, and this is referred to as “same view advantage” (for a review see Liu & Chaudhuri, 2002).

1.3.2.2.1 Different view advantage

A different view advantage is apparent if the ¾ view shown at the study or test produces allows better recognition performance than other views (such as full-face or profile). This effect was found predominantly in studies using a recognition paradigm, where subjects were shown a target face in the learning period and then were required to identify the target among distracters in the test period (Patterson & Baddeley, 1977; Woodhead, Baddeley, & Simmons, 1979; Krouse, 1981; Baddeley & Woodhead, 1983; Logie, Baddeley, & Woodhead, 1987; Schyns & Bülthoff, 1994; Valentin, Abdi, & Edelman, 1997, O`Toole, Edelman, & Bülthoff, 1998). However, there are also recognition studies showing no advantage for the ¾ view (Davies, Ellis, & Shepherd, 1978; Hill, Schyns, & Akamatsu, 1997; Liu &

Chaudhuri, 1998). A few studies have also used a sequential matching paradigm, where subject were required to judge whether two consequently displayed faces were of the same person, reporting mixed results (Troje & Bülthoff, 1996; Bruce et al., 1987; Liu, Collin, Burton, & Chauduri, 1999).

The mixed reports prompt the question towards the source of this inconsistency. Although mixed results are often attributed to methodological differences, a more important factor may arise from two mutually inconsistent definitions of a different view advantage. One definition depends on the advantage produced by angular rotation. The other focuses on the potential advantage due to intrinsic properties of the ¾ view.

The ¾ view advantage based on angular rotation is well illustrated in a study by Patterson and Baddeley (1977). They examined recognition performance for faces initially trained in full-face views, but subsequently tested in ¾ and profile view. The pose change was also combined with disguise (such as a change in a wig or beards or both). Results showed that targets were recognized better in ¾ view than in profile view, regardless of whether there was a disguise or not. However, the ¾ view advantage over the profile view found in this study is most likely based on the fact that the target faces (full-face) underwent a greater angular rotation in depth for the profile (90° rotation) than for the ¾ view (45° rotation). This does not rule out the possibility that the advantage is partially contributed by some inherent characteristics of the ¾ view, but it is impossible to estimate this based on the paradigm used.

There are examples of other studies where the ¾ advantage also depends on the effect of angular rotation between learning and test views (Baddeley & Woodhead, 1983; Logie et al., 1987; Schyns & Bülthoff, 1994; Troje & Bülthoff, 1996; Valentin et al., 1997). For example, Baddeley and Woodhead (1983) presented the target in either frontal, ¾, or profile view in the learning period, and compared recognition under each of these conditions in the test period. The conditions that implied 45° rotation (frontal to ¾ view, profile to ¾ view and vice versa) yielded the best performance followed by the conditions implying 90° rotation (frontal to profile and vice versa). A separate control experiment in their study showed that faces trained in ¾ view were recognized better when the overall performance for the three testing views (full face, ¾ and profile) was combined. The analysis revealed that the frontal and profile views suffered from angular rotation between learning and test view about 90°. This suggests that the ¾ view advantage would have likely disappeared if the amount of angular rotation had been equated, so the total angular rotation amount was the confounding variable.

1.3.2.2.2 Same view advantage

As mentioned earlier, a same view advantage refers to the better identification of faces in ¾ view than other views when the training and test view are identical. As for the different view advantage mixed results were reported in studies that have tested the same view advantage.

An early positive finding within this category was reported by Fagan (1979), who tested face recognition in 7-month-old babies using two pairs of faces. The results showed that the babies were able to discriminate the new faces from the old ones. Furthermore, a superior performance was found for faces presented in ¾ view than in full-face or profile view when the pair with similar faces was used.

Bruce et al. (1987) found that adult subjects identified faces faster when the face images were presented in ¾ view than when they were presented in full-face or profile views. They did not find any difference in their accuracy data, whereas the speed effect was present only for unfamiliar, but not for familiar faces. Unlike Bruce et al. (1987) who used a sequential matching task, Valentin et al. (1997) and O`Toole et al. (1998) compared the effect of the three views in a recognition task.

Both studies reported a ¾ view advantage in comparison to the full-face and profile views.

However, there are studies that did not find a ¾ view advantage (Logie et al., 1987; Harries, Perret, & Lavender, 1991; Hill et al., 1997; Newell et al., 1999). For example, Hill et al. (1997) showed that performance for the ¾ view was nearly identical to that of full-faces view. A surprising finding was that performance for the profile view was better than for full-faces and ¾ views. Newell et al. (1999) found a significantly poorer recognition for ¾ view than for the full-faces view.

In sum, compelling evidence for a same view advantage is as difficult to find as for a different view advantage.

1.4 Timing of perception

1.4.1 Time course of object and face perception

A number of physiological and behavioral studies demonstrated that object categorization and recognition is performed rapidly (Fabre-Thorpe, Richard, & Thorpe, 1998; Intraub, 1999, Breitmeyer, 1984, Grill-Spector, Kushnir, Hendler, & Malach, 2000). Fabre-Thorpe et al. (1998) investigated how much time observers need to judge whether a visual scene contains an animal or not. They showed that a correct motor response can be generated by humans as quick as 235 ms after the scene was presented (Fabre-Thorpe et. al., 1998). Rapid serial visual presentation (RSVP) techniques (Intraub, 1999) and other masking paradigms (Breitmeyer, 1984) reveal behavioral selectivity to images presented for 100 ms or less. Human functional magnetic resonance imaging (fMRI) studies have shown that activation in object recognition areas correlated with masked images arises with presentations as brief as 40 ms (Grill-Spector et. al, 2000).

In contrast to previous studies suggesting that the early presentation-locked component of neural activity is correlated with recognition, results of a study by Johson and Olshausen (2003) implied that the neural signatures of recognition have a substantially later and variable time of onset. They found two types of components in the ERP recorded during categorization of natural images. One is an early presentation-locked signal arising after about 135 ms that is present when there are low-level feature differences between images. The other is a later, recognition related component arising after about 150-300 ms.

In a recent behavioral study Grill-Spector and Kanwisher (2005) varied exposure duration of natural images and measured subjects` performance in three different tasks: object detection, object categorization and within-category identification. In the object detection task subjects were asked to decide whether or not a gray-scale photograph contained an object. In the object categorization task, subjects were asked to categorize the object in the picture at the basic level (e.g. car, hose, flower). In the within-category identification task, subjects had to discriminate exemplars of a particular subordinate-level category (e.g. German shepherd dog) from the other members of the category (other dogs). They presented each photograph briefly at one of several exposure durations (17, 33, 50, 68, or 168 ms) and instantaneously masked, measuring percent of correct responses and reaction time. The result of their study suggests that subject did not required more processing time for object categorization than for object detection. Instead, as soon as subject could detect an object at all, he/she already knew its category. However, comparable performance on the identification task required substantially longer exposure duration and more processing time than was required for either detection or categorization. On average, 65 ms more were necessary for identification than for categorization (see Figure 1.4).

Less clear so far is the time scale at which processes related to different aspects of face perception operate. Face-responsive event related potential (ERP) activity was initially described in form of a vertex positive potential (VPP) with

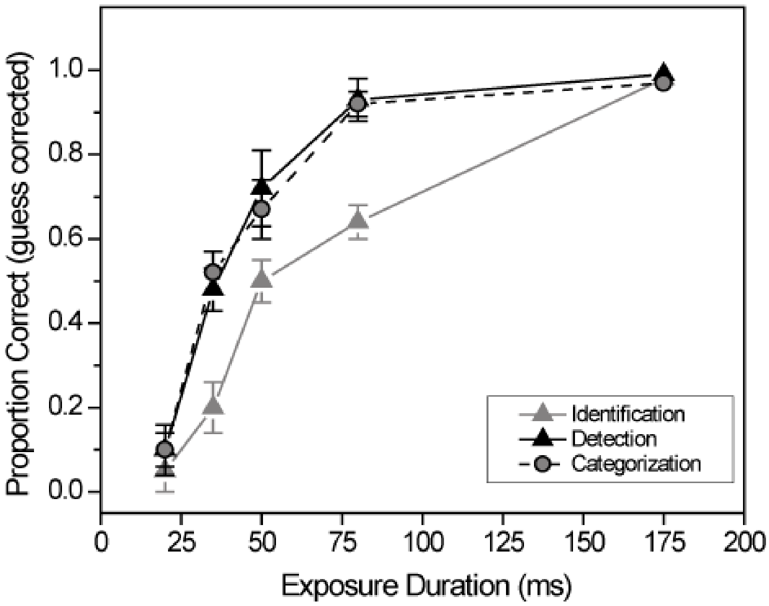

Figure 1.4: Results of the combined detection/categorization/identification task of Grill-Spector and Kanwisher (2005). The data are reproduced from their Figure 2. The y-axis indicates accuracy in a proportion correct measure corrected for guessing. Error bars indicate standard errors of the means.

reference to the ears that was observed 150-200 ms after stimulus presentation and had larger amplitude and shorter latency for faces as compared to other complex or simple visual stimuli - fragmentary figures, schematic objects, animals (Jeffreys, 1989). A large negative ERP component with a latency of 170 ms (N170) is taken to represent a face-specific response generated by brain mechanisms involved in face processing. It is suggested that N170 is linked to late stages of structural encoding, related to the categorization of stimuli as faces and to the recognition of individual faces, whereas a response occurring about 100 ms after stimulus onset is correlated only with successful categorization (Eimer, 2000; Liu, Harris, & Kanwisher, 2002). These data suggest that face processing proceeds trough two stages: an initial stage of face categorization, and a later stage at which the identity of individual faces is extracted (Liu et al, 2002).

The study of Eger, Jedynak, Iwaki, and Skrandies (2003) provides evidence about the time course of perceptual processes related to facial expression, indicating that early VEP (visual evoked potential) components occurring at 80-90 ms are sensitive to emotional content. Single unit recordings in macaque superior temporal sulcus show that face-selective activity can be elicited by masked images presented for as little as 14 ms (Keysers, Xiao, Foldiak, & Perrett,. 2001).

The time course of face processing was the subject of behavioral studies, too. Yin (1970) found a disproportional effect of brief presentation which is similar to the disproportional effect of inversion. Reducing exposure duration from 5 sec to 100 ms caused a greater impairment in recognition of faces than in recognition of bridges or stick figures. Lindsay, Jack, and Christian (1991) found in a study with faces of other ethnic groups that Caucasian participants were able to acquire and retain more useful information about Caucasian faces than about African American faces, when very brief (120 ms) exposure duration were used. Lindsay et al. (1991) suppose that no such difference would be obtained with longer exposure duration (i.e. the effect might be due to differences in the rate at which useful information is acquired). Hole (1994) used vertically and horizontally chimeric faces, which were presented for 80 ms or 2 sec. He suggested that long exposure time allows feature by feature comparison, while short exposure time forces the participant to use rather configural strategies to process faces as wholes (Hole, 1994).

In a recent study Lehky (2000) showed that the smallest change in a face can be discriminated very fast. He used synthetic faces created by morphing together pairs of source faces. For stimulus exposure duration longer than 100 ms subjects were able to discriminate two faces when their difference reaches about 7% of the average face difference. Longer exposure duration (up to 1000 ms) did not result in a better discrimination performance. However, below 100 ms the discriminability of faces deteriorates abruptly. At 53 ms faces where discriminable when their difference reached 17% of the average face difference. The results of this study are shown in Figure 1.5.

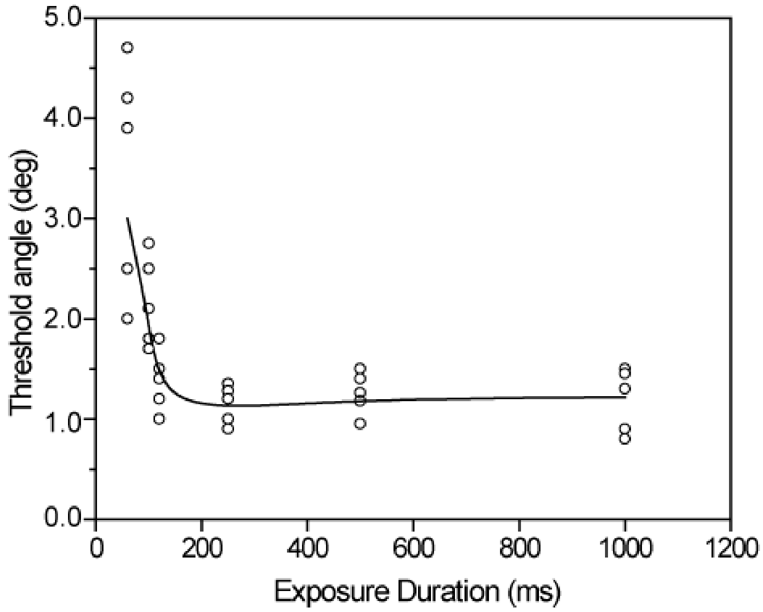

Figure 1.5: Face-discrimination thresholds as a function of stimulus duration for one of three subjects (CJ). The y-axis is the vector angle difference between sample face and threshold face, and the x-axis is the exposure duration of the test face. At each duration, circles show thresholds for the eight face discriminations and the smooth line indicates a spline approximation to the means (Lehky, 2007, Figure 3, page 850).

Studying the early stages of face recognition and the role of featural and holistic face information, Carbon and Leder (2005) used thatcherised faces in a speed identification task with limited exposure duration times (26 and 200 ms). “Thatcherisaton” of faces originates from turning the eyes and the mouth region upside down (Thompson, 1980). In upright orientation, thatcherised faces are perceived as very grotesque faces. However, when inverted, the alienation is hardly detectable (Bartlett & Searcy, 1993; Lewis & Johnston, 1997) (see Figure 1.6 for an example of original and Thatcher faces).

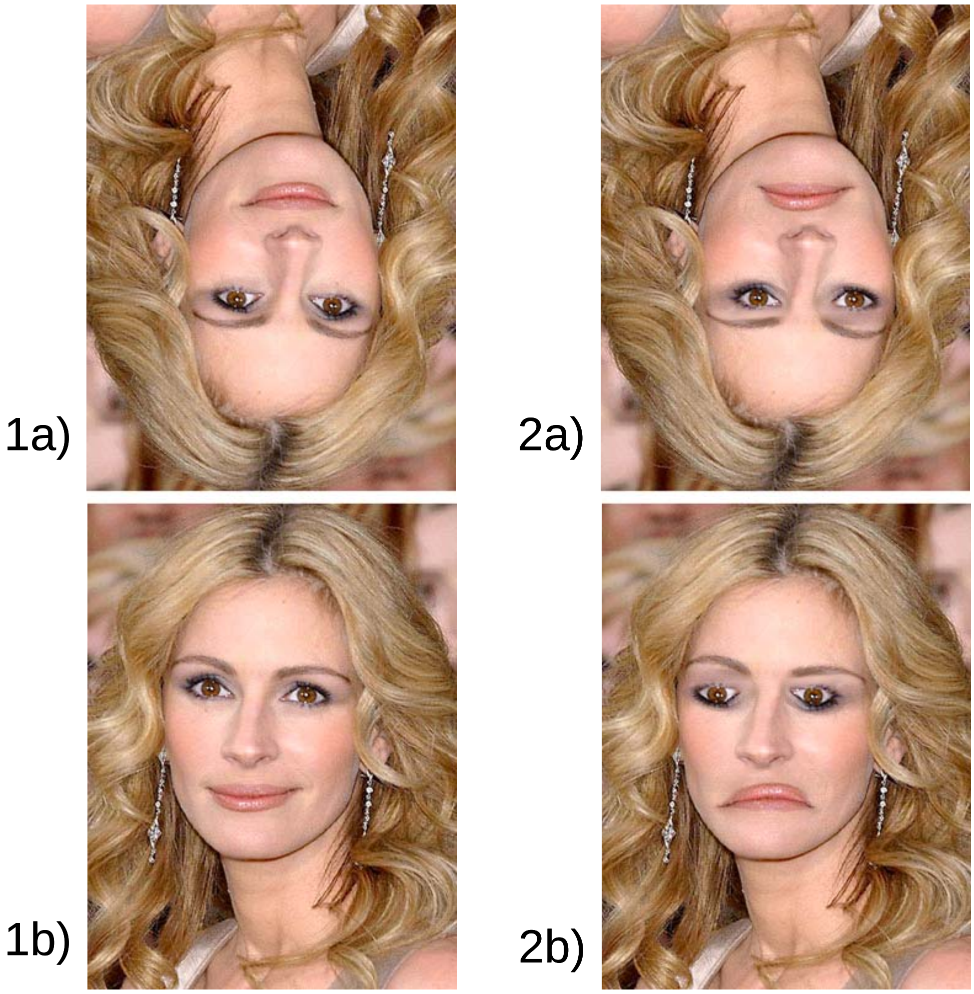

Figure 1.6: Example of original (left) and Thatcher faces (right) in inverted and upright orientation. The Theatcher face (right) is perceived as grotesque only in upright orientation (2b), but not when turned upside down. Similar stimuli were created and used in Experiment 1 and 3 in the study conducted by Carbon and Leder (2005).

By using original and Thatcher faces featural and holistic information can be dissociated. Carbon and Leder (2005) reason that if early processing of features is beneficial for the identification of faces, then inverted Thatcher faces should be processed faster than inverted originals at brief exposure durations of about 26 ms, since no mental rotation of the features (eyes and mouth) is required for comparison. Alternatively, if features are processed at later stages of processing an advantage is expected only at longer exposure duration (200 ms). On the other hand, if holistic processes are beneficial for face processing, then inverted original faces are expected to be identified faster than inverted Thatcher faces due to the holistic coherence of the original. Indeed, the experiment clearly showed that at 26 ms of exposure duration inverted Thatcher faces were processed faster than inverted original faces. At longer exposure durations of 200 ms performance was better with inverted original faces than with inverted Thatcher faces. The authors interpreted this interaction as evidence for a temporal precedence of featural information over configural information. At brief timings only featural information is available, allowing the observer to perform just feature based comparisons. Later, configural information adds, allowing the observer to exploit relations among features in his/her judgements. With this interpretation the authors assume that inversion does not totally preclude configural information to become effective, but assume that this kind of information gradually evolves with time, and is used also in inverted faces. With more time the observer is able to see that, despite inversion, there is something wrong with a Thatcher face.

2. Research Questions and Experimental Rationale

In three studies the relationship of featural and configural face processing was investigated by exploring the time course of processing depending on inversion and viewpoint as two relevant diagnostic factors in face perception. Considering facial features, a distinction has been made between external (e.g. hair, ears, facial and head outline) and internal (e.g. eyes, eyebrows, nose and mouth) features. Present findings suggest that internal and external features might play different roles in face identification. These different roles have widely been discussed in studies dedicated to the different processing schemes found for familiar and unfamiliar faces (Ellis, Shepard, & Davies, 1979; De Haan & Hay, 1986; Young, Hay, McWeeny, Flude, & Ellis, 1985; Hines, Jordan-Brown, & Juzwin, 1987; Bruce et al., 1999; Hancock, Bruce, & Burton, 2000; Jarudi & Sinha, 2003; Frowd, Bruce, McIntyre, Hancock, 2007), indicating that internal features are most important for handling familiar faces while external features are focused during processing of unfamiliar faces. However, in all these studies internal and external features were presented in isolation, without facial context (see Figure 2.1). Therefore, the conclusion about the close relation of familiarity and feature class preference is not proven to hold in the natural facial context, where internal and external features interact, as they do in everyday vision.

Although it is interesting to learn about processing of internal and external features in isolation, we like to know how a variation of internal features affects face recognition in the presence of (constant) external feature context, and vice versa. If manipulated within facial context the contribution of external and internal features for the processing of whole faces can be examined, and the role of each feature class can be assessed in its natural context (see Figure 2.2).

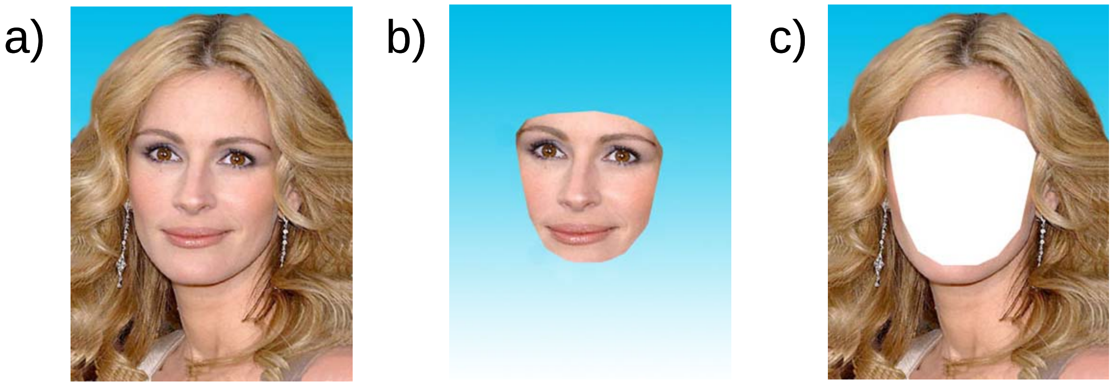

Figure 2.1: An example of a whole face (a), its isolated internal (b) and external(c) features, as they have been used in previous studies.

With a given facial context we aim to address part-to-whole processes by studying how external and internal features account for the perception of whole faces, but with the weight layed upon either feature class, as controlled by instruction. Arranging tasks such that they can be solved only by focusing the feature of interest, it is possible to study how the feature class under inquiry is modulated by external factors such as viewpoint, inversion, and exposure duration – but always in the context of an entire and intact face. Employing completely balanced stimulus material with respect to its composition of internal and external features it is possible to let subjects do the same type of matching task with different classes of features but the same set of facial stimuli, which appear as intact faces and therefore as natural face stimuli.

This makes it possible to selectively tap the processing paths involved in handling internal and external feature information during processing of natural face stimuli, and to study how these paths are affected by external factors. Differential results for internal and external feature processing with respect to external influence factors will make it possible to characterise these two distinct processing paths, and to draw conclusions about one basic question of face perception - whether face processing is featural or configural - given these two modes are supported by differential findings about the influence of viewpoint and inversion.

Employing the rationale outlined above the effects of inversion, viewpoint and timing are addressed with a series of three studies. In the first two studies the effects of inversion and viewpoint were particularly focused, while the last study was particularly dedicated to measuring the timing of processing internal and external features. If there are different processing paths for internal and external features, then both should be differently affected by viewpoint and inversion. Existing evidence suggests that internal and external features differentially tap separate functional entities of the cognitive system: presumably, internal features affect the face recognition system while external features are processed by the

Figure 2.2: An example of the stimuli we have used in our studies: internal and external features of original faces (a) have been combined resulting in two new faces (b).

object recognition system (Moscovitch & Moscovitch, 2000). Furthermore, it has been proposed that the face recognition system primarily relies on configural information, as provided by internal features in the usual (upright) orientation, whereas the object recognition system operates with featural and part-based information, and is involved in the processing of inverted faces (Moscovitch et al., 1997; Moscovitch & Moscovitch, 2000). By exploring how inversion and viewpoint selectively affect internal and external feature processing within an intact facial environment we are able to test not only whether both feature classes are differentially processed, we can also account for the contribution of each of the two hypothetical systems for the given task.

An important aim of the studies was to estimate the time course of feature and whole face processing, which is a new contribution to the field. To end this we varied exposure durations of faces with effective masking, controlling the available amount of information which can be extracted and used for completing a matching task. This aimed at determining critical exposure durations for different degrees of performance: maximum performance, half the maximum performance, and, the conventional level of 75% correctness in the experimental condition. As discussed in section 1.4.1, the time necessary to perform a task at a given level of accuracy reveals involvement of additional processing stages, and is a relatively sensitive and reliable measure apt to indicate whether the same or different processing paths are used for different stimulus material or task instructions. Although there are some neurological and behavioural data suggesting that a certain amount of time is related with different aspects of face processing (Jeffreys, 1989; Liu et al, 2002; Lehky, 2000; Carbon & Leder, 2005), fundamental questions remain hitherto to be addressed. First, it is unclear whether whole faces are processed faster than their features, which points to the question of the speed of configural and featural modes of processing. Second, related to the former question, there are no measurements at the time allowing to judge whether external and internal features are processed at the same time scale. Combining the effects of viewpoint and inversion with the timing aspect gives valuable hints at the existence of distinct modes of processing in face perception. If it turns out that internal features are subject to inversion and change of viewpoint, and their accurate processing requires a reasonable amount of time, while processing of external features is fast and proves to be relatively invariant across changes in viewpoint and orientation, then a relevant contribution to the question about the two distinct modes of face processing is at hand.

3. The Time Course of Processing External and Internal Features of Unfamiliar Faces

3.1 Summary

The time course of processing internal and external facial features was studied in a sequential face matching task, where first a target face was presented, followed by a test face. The exposure duration of the test face was varied systematically (90 ms, 120 ms, 150 ms, and self-paced). In three tasks, participants were instructed to match either the whole face, only external features, or only internal features of the target and test face. Taken together, the results in all three tasks provide evidence for very fast matching processes. For upright faces, maximal performance was achieved at 90 ms exposure duration and longer exposure durations (120 ms, 150 ms, self-paced) did not improve accuracy. For inverted whole faces, reduced exposure duration resulted in an increase of matching errors, suggesting that below 150 ms of exposure duration, inverted faces cannot be matched reliably. When matching selected facial features only, no such inversion effect was found. Our data challenges previous claims that external features are matched faster than internal: no difference of time course was found between external and internal features. However, external features were matched more accurately.

3.2 Introduction

Humans are supposed to recognize faces with astonishing ease and speed (Seeck et al., 1997). However, less is known about the time scale and operations occurring in the early stages of face processing. Neuropsychological studies have shown that visual ERP (event related potential) for facial stimuli is elicited within less than 200 ms (Jeffreys, 1989; Eger et al., 2003; Keysers et al., 2001). The time course of face processing has also been the subject of behavioral studies. Yin (1970) found a disproportionate reduction of recognition performance under brief presentation (100 ms) which is similar to the disproportionate effect of inversion in faces. Hole (1994) used vertical and horizontal chimeric faces, which were presented for 80 ms or 2 sec. He suggested that long exposure duration allows feature by feature comparison, while short exposure duration forces participants to use configural strategies to process the face as a whole (Hole, 1994).

In our study, we used behavioral methods to investigate the sequence of processes involved in matching unfamiliar faces. Our focus lay in the processing of external and internal features in the whole facial context. External facial features refer to the hair, ears, head and face outline, whereas internal features refer to the eyes, eyebrows, nose and mouth. Three hypotheses were tested:

1) The first hypothesis deals with the effect of orientation. The claims regarding the inversion effect for isolated internal or external features are ambiguous. It has been argued that since inversion impairs configural (spatial relationship between features) but not featural processing (e.g. hair, face outline, eyes, mouth, nose etc.), recognition of facial features in inverted faces should not be impaired (for a review see Valentine, 1988; Schwaninger, Carbon, & Leder, 2003). However, the inversion effect was found for isolated facial features (Barton et al., 2001; Rakover & Teucher, 1997) as well as for the external and internal features (Nachson & Shechory, 2002; Moscovitch & Moscovitch, 2000). In all of these studies the external and internal features were presented isolated, without the

whole facial context. Consequently, the conclusions gained from those studies can not be applied to the processing of the external and internal features within the whole facial context. The contribution of the facial features to face perception can be studied only in the context of whole faces, where not only featural but also configural or holistical information is available (Nachson, Moscovitch, & Umilta, 1995). According to Nachson et al. (1995), upright faces are processed on the basis of their configurations in the sense that the various features interact to produce an integrated facial stimulus. They, however, did not investigate if there is any configural integration of the facial features if stimuli are presented inverted. Moreover, black and white line drawings were used as stimuli, which are not a realistic representation of a face. The aim of our study was to investigate whether orientation impairs matching of internal or external features when they are shown in a facial context of a photographic image. If the configural relations are weaker or completely lost in inverted orientation we may expect no effect of inversion for matching of only internal or only external facial features.

We were also interested in how exposure duration affects matching of upright and inverted faces. There are two reasons why such an effect could be expected: First, there is evidence that reduction of ERP response for identical inverted faces is smaller and delayed by 30 ms in comparison with identical upright faces (Jacques, d'Arripe, & Rossion, 2007). Second, if inverted faces are processed by a serial part-by-part strategy, which is also time-consuming, we might expect an interaction between exposure duration and inversion. Barton and colleagues (2001) have found that inversion effect was reduced with prolonged exposure duration, whereas the performance for upright faces stayed unaffected. However, they used rather long exposure durations (1 sec, 2 sec, 3 sec and self-paced) which do not have direct implication for early face processing (Barton et al., 2001).

2) An interaction between feature configuration and the exposure duration was addressed in this study too. Matching performance is expected to be strongly impaired by the features of the faces being compared. A match in only internal or

external features of two faces is found to be more difficult than the matching of wholly identical or wholly different faces (Nachson et al., 1995). The matching of faces sharing only external or internal features ("same external", SE; "same internal", SI) requires more part-by-part comparisons, whereas two faces, which are either identical (ID) or have no common features at all ("different", DF) could be processed more holistically. This seems to be even more so since the reaction times are longer for the condition SE and SI than for ID and DF (Nachson et al., 1995). Consequently, the matching of two faces sharing only external (SE) or internal (SI) features is expected to be more impaired by short exposure duration than matching of identical or different faces.

3) The studies of the microgenesis of perception led to the hypothesis that global aspects of objects are processed faster than their details (Love, Rouder, & Wisniewski, 1999; Hubner, 1997; Hoeger, 1997). In face recognition, most of the studies have revealed faster and more accurate matching of external than internal features (de Haan & Hay, 1986; Nachson et al., 1995; Young et al., 1985). Internal features contain featural (eyes, eyebrows, nose, mouth) as well as configural information (distances between those features), while external features contain more global information. According to this assumption, internal features are expected be more impaired by short exposure duration than external features, because the former contain more information and have a more detailed structure. External features refer to the global aspects of the face and therefore are expected to be less impaired by short exposure duration.

To test our hypotheses we measured behavioral performance in three different matching tasks: matching of whole faces (Task 1) and matching of only external (Task 2) or only internal (Task 3) features. A sequential matching task was used, with constant exposure duration for the target faces (1500 ms), and four different exposure durations for the test faces (90 ms, 120 ms, 150 ms and self-paced). The short exposure durations were varied within the range of the ERP time course for faces (Jeffreys, 1989; Eger et al., 2003; Keysers, et al., 2001).

Presentation of the target and test faces were separated by 1000 ms of an interstimulus interval (ISI) for two reasons. Firstly, we aimed to avoid a participant's use of traces in iconic memory, which facilitates complete pictorial matching instead of face matching. It has been showed that representation in iconic memory lasts up to 200-300 ms and decays after that (Schilller, 1968; Rolls & Tovee, 1994; Kovacs, Vogels, & Orban, 1995; Keysers, Xiao, Foldiak, & Perret, 2005; Martens, Schweinberger, Kiefer, & Burton, 2006). After 1000 ms no traces of the first presented, target face is expected to be present in iconic memory. Secondly, self-paced exposure duration, where the second face was exposed until the response, was performed to compare performance between short and long exposure durations. A mask can not be used in this test condition. To keep test conditions comparable, we did not use a mask for the short exposure durations (90 ms, 120 ms, 150 ms). In spite of lacking the mask we believe that the exposure duration should still have an effect on the general performance. As it is well established, information present in iconic memory is pre-categorical and if to be processed an item must be represented in post-categorical form (for a review see Coltheart, 1983). Moreover, the transfer to post-categorical stage depends strongly on the on the number of items and on the capacity limits of working memory (Coltheart, 1977, 1983; Fuster & Jervey, 1981). Our tasks are rather complex, requiring comparisons on a part-to-whole basis relying strongly on the memory for both target and test faces. That is why we expect that viewing exposure durations in combination with the long ISI between the two faces should affect matching performance, even if the visual persistence was not controlled by a mask. To ensure that the pictorial matching and the lack of the mask did not influence general performance we conducted a control experiment where the whole identity of the two faces was matched within the 150 ms exposure durations. In the control experiment the second presented picture of the face was 17 % smaller than the first and a mask was presented for 100 ms directly after the second face. Matching of the whole identity between faces is supposed to be easier than the part-based matching of isolated features (Nachson et al, 1995). Therefore, if the lack of a mask does not affect matching of the whole face in Task 1, then we can argue that it would have an even smaller influence on the other two part-based matching tasks (matching of external or internal features).

3.3 Method

3.3.1 Participants

A total of thirty-eight students (age 20 - 35 years), participated in the three experiments introduced above: ten in Task 1 (9 females), fourteen in Task 2 (10 females), and fourteen in Task 3 (11 females). In the control experiment, 10 students (7 females) were all unaware of the purpose of the experiment and none had participated in the previous three experiments. All participants had normal or corrected to normal vision.

3.3.2 Stimuli

Full-color frontal view photographs of 12 male faces were captured in a photo studio under controlled lighting conditions using the same background. None of the faces was wearing glasses and jewelry or had a beard. The haircut for all faces was kept comparably similar with short hairstyle. Four faces were chosen as the target faces and were paired with the appropriate twelve test faces. For each of the 4 target faces there were 12 test faces respectively. Three of the test faces were identical with the target face, three were sharing the same internal features as the target faces, three were with the same external features and three of the test faces were completely different from the test face.

Figure 3.1: Sample of a target face and three test faces.

A total of possible pairs were forty-eight. The test faces were assigned to four conditions. In the identical (ID) condition, the same face was used as test and target face. In the same external (SE) condition, both target and test faces shared the same external features, while internal features differed. In the same internal (SI) condition, only the internal features of the target and test faces were the same. In the different (DF) condition, target and test faces depicted two different persons with no common features at all. Figure 3.1 shows an example of a target and three corresponding test faces. The test faces for condition SE and SI were prepared using Adobe Photoshop 9. The internal features were cut out with comparable tracing lines and placed on the second (template) face, based on the position of the internal features. All stimuli were 14.11 cm long and 10.58 cm wide with the resolution 72 pixels/inch. Images were presented on a color 17'' CRT monitor. Screen resolution was set to 1024 x 768 pixels with the refresh rate of 60 Hz. The viewing distance was approximately 60 cm.

3.3.3 Procedure

The participants were tested individually in the experimental room. All participants first completed a short practice to become familiar with the task. All were tested in two sessions, once with upright and once with inverted stimuli (both target and test faces), and were randomly assigned to one of the two procedures (upright first versus inverted first). Between two sessions, there was a break of seven days in order to avoid a learning effect. One session took about 30 minutes and contained 192 trails. Each trial started with a target face that was displayed for 1500 ms, and after a blank ISI of 1000 ms, a test face appeared. Images were presented centered using SuperLab 2.0. Participants were asked to respond as quickly and accurately as possible whether the target and the test faces were the same or different by pressing the left or right mouse button with their preferred hand. The assignment of answers (*same*/*different*) to the left or right mouse button was counterbalanced across participants. In Task 1, *same* was defined as whole congruency between target and test faces, including both internal and external features. In Task 2, *same* was defined as congruency of external features only, and in Task 3 as congruency of internal features only. Participants completed either Task 1, or 2, or 3. Each test face was shown four times, once for 90 ms, 120 ms, 150 ms, and once until the participants' answer occur (self-paced condition). Four blocks with different exposure durations were randomized across participants, while condition (ID, SI, SE, DF) was randomized across trials. In the control experiment matching of whole faces was required (as in the Task 1), whereas the exposure duration was constant at 150 ms and the test face was 17% smaller than the target face.

3. 4 Results

The reaction times of correct responses and the number of errors were recorded for each trial. The data of the three tasks were combined and subjected to a 4-factor analysis of variance (ANOVA) with orientation (upright vs. inverted), exposure durations (90 ms, 120 ms, 150 ms and self-paced) and conditions (ID, SE, SI, DF) as within-participants factors and task (same = wholly identical, same = identical internal features only, same = identical external features only) as between-participant factor. Bonferroni correction was conducted for multiple comparisons. Additional analyses were performed where necessary as reported below.

3.4.1 Average Error

Results for repeated ANOVA for the mean matching errors are shown in the Table 3.1.

Source	SS	df	MS	F	p
TASK (A)	1.85027	2	0.92513	15.2445	0.000017
Error	2.12402	35	0.06069		
ORIENTATION (B)	0.47761	1	0.47761	7.1732	0.011194
A x B	0.31861	2	0.15931	2.3926	0.106181
Error	2.33036	35	0.06658		
EXP. DURATION (C)	0.15054	3	0.05018	5.4118	0.001679
A x C	0.02774	6	0.00462	0.4986	0.808174
Error	0.97357	105	0.00927		
CONDITION (D)	2.31503	3	0.77168	20.0912	0.000000
A x D	5.81416	6	0.96903	25.2294	0.000000
Error	4.03291	105	0.03841		
B x C	0.10931	3	0.03644	2.8212	0.042472
B x C x A	0.05278	6	0.00880	0.6811	0.665221
Error	1.35616	105	0.01292		
B x D	0.33050	3	0.11017	3.5904	0.016165
B x D x A	0.23306	6	0.03884	1.2659	0.279444
Error	3.22174	105	0.03068		
C x D	0.13954	9	0.01550	1.4266	0.175520
C x D x A	0.19895	18	0.01105	1.0169	0.439925
Error	3.42359	315	0.01087		
B x C x D	0.08150	9	0.00906	0.9319	0.497398
2*3*4*1	0.17507	18	0.00973	1.0009	0.458233
Error	3.06088	315	0.00972		

Table 3.1: Repeated ANOVA for mean matching errors. The Task (matching of whole faces, external or internal features) was a between-participants variable, and Orientation, Exposure Duration and Conditions were within-participants variables.

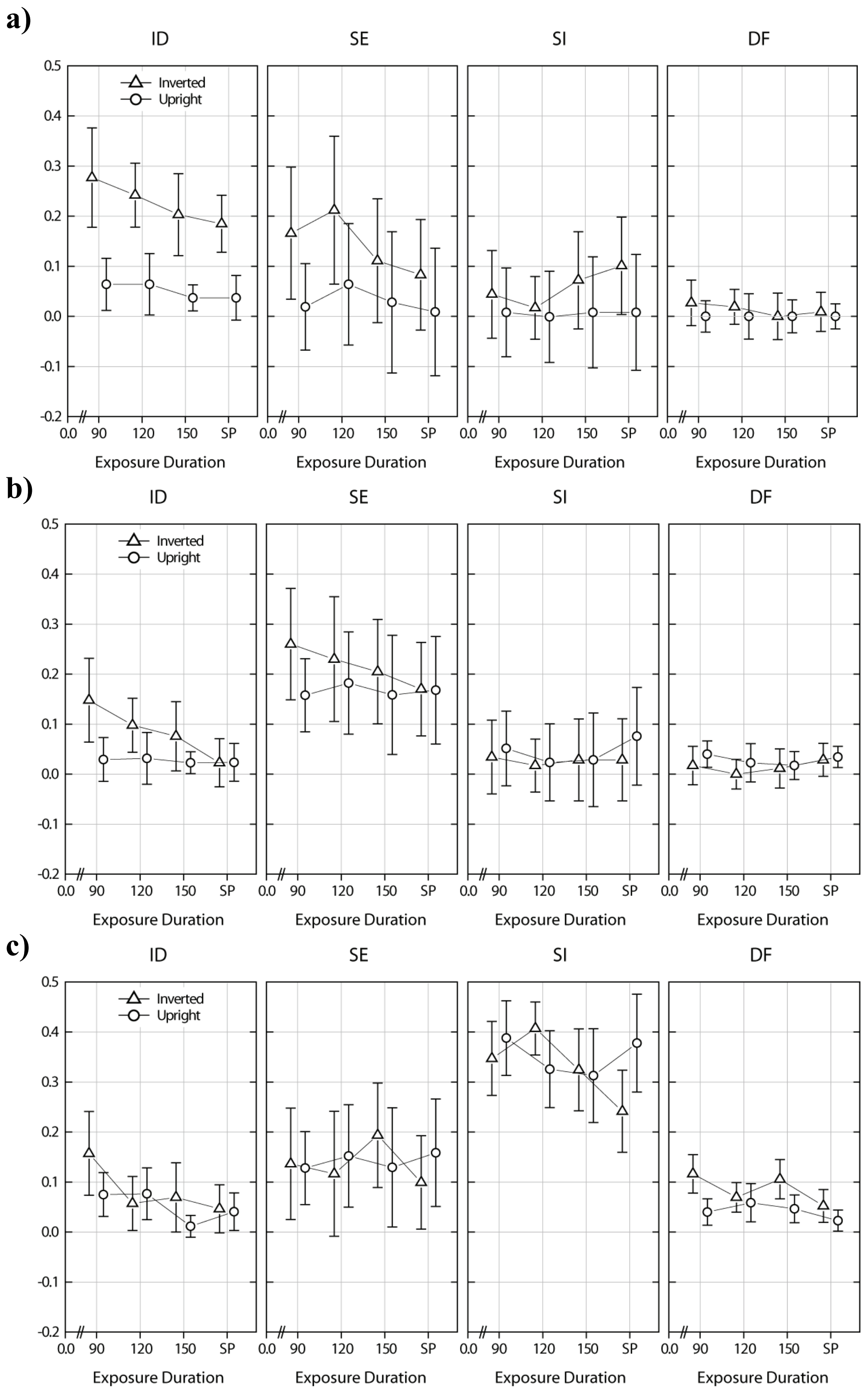

Figure 3.2: Average matching errors for a) Task 1 – matching of whole faces, b) Task 2 – matching of external features and c) Task 3 – matching of internal features in upright and inverted orientation. Average number of errors are plotted against the four exposure durations and broken up by conditions (ID, SE, SI, DF).

The analysis of matching errors revealed significant differences between the three tasks, whereas pairwise comparisons showed that participants were less accurate in Task 3 (matching of internal features) than in Task 1 (matching of whole faces) and Task 2 (matching of external features) (both $p < .001$). There was no significant difference in matching errors between Task 1 and Task 2. Average matching errors for all three tasks are depicted in Figure 3.2.

The main effect of condition was also significant as well as an interaction between task and condition. For identical-trails (ID), the average matching error was significantly higher than for matching of SE-, SI- or DF-trails in Task 1 (same = identical faces). In the Task 2 (same = same external features) matching of the SE-trails was related with significantly more matching errors than conditions ID, SI and DF. In Task 3 (same = same internal features) the highest matching error was found for SI-trails in comparison with the other three conditions (all $p < .001$).

Overall, the performance for upright faces was more accurate than for inverted ones and there was a significant interaction between orientation and condition. The simple effect of orientation was significant only for ID-trials $F(1, 37) = 21.39$; $p < .001$, whereas there were no orientation effect for SE-trials ($F(1, 37) = 1.64$, $p = .201$), SI-trials ($F(1, 37) = 0$; $p = .98$), and DF-trials ($F(1, 37) = 2.4$; $p = .1.2$). A separate analysis for each of the tasks showed that the inversion effect reached statistical significance only in Task 1: $F(1, 9) = 23.11$, $p < .001$, where whole faces had to be matched. In Task 2 (matching of external features) and Task 3 (matching of internal features), there was no significant inversion effect (Task 2: $F(1, 13) = 1.20$, $p = .29$; Task 3: $F(1, 13) = .13$, $p = .72$).

The main effect of exposure duration was significant in the sense that matching errors decreased with the increased exposure durations. The pairwise comparisons revealed a significant difference in matching errors between 90 ms and self-paced condition. Moreover, there was a significant interaction between orientation and exposure duration. The simple effect of exposure duration was found for inverted but not for upright stimuli: $F(3, 105) = 7.48$, $p < .001$. Pairwise comparisons showed that there was no significant difference (p = .88) in matching errors between self-paced in comparison with the other three exposure durations (90 ms, 120 ms and 150 ms) if stimuli are presented upright. The same pairwise comparisons for inverted presented stimuli showed a significant effect (p < .001) of exposure durations, with a significant difference appearing below 150 ms.

3.4.2 Reaction Times

The results of the repeated ANOVA are shown in the Table 3.2, whereas the mean reactions times are depicted in the Figure 3.3.

ANOVA revealed no differences in reaction times between the three tasks suggesting that whole faces as well as internal and external facial features were processed equally fast. Overall, reaction times were faster for upright than for inverted presentations. Although an interaction between orientation and condition for reaction times approximated statistical significance, the analysis of simple effects showed the same pattern as found for accuracy: The inversion effect for ID-trials was significant, $F(1,37) = 8.7$, $p < .001$, while for SE-trials, SI-trials and DF-trials, there was no significant effect of orientation. Separate analyses for each of the three tasks revealed significant effects of orientation only in Task 1 (matching of whole faces): $F(1, 9) = 23.11$, $p < .001$. There was no orientation effect in Task 2 (matching of external features): $F(1, 13) = .12$, $p = .74$, and Task 3 (matching of internal features), $F(1, 13) = 1.15$, $p = .30$.

The main effect of condition and an interaction between task and condition were significant. As for accuracy pairwise comparisons showed that the reaction time was the longest for ID-trials in the Task 1 (matching of the whole faces), SE-trials in Task 2 (matching of external features) and SI-trials in Task 3 (matching of the internal features) (all p < .001).

Source	SS	df	MS	F	p
TASK (A)	187903	2	93951	0.0923	0.912052
Error	35625717	35	1017878		
ORIENTATION (B)	3346384	1	3346384	4.3003	0.045528
A x B	3186297	2	1593148	2.0473	0.144263
Error	27236224	35	778178		
EXP. DURATION (C)	7242625	3	2414208	17.4781	0.000000
A x C	877752	6	146292	1.0591	0.391928
Error	14503406	105	138128		
CONDITION (D)	4630302	3	1543434	17.3648	0.000000
A x D	10579306	6	1763218	19.8376	0.000000
Error	9332681	105	88883		
B x C	56747	3	18916	0.1831	0.907701
B x C x A	611375	6	101896	0.9861	0.438553
Error	10849868	105	103332		
B x D	610201	3	203400	2.5240	0.061656
B x D x A	582862	6	97144	1.2054	0.309347
Error	8461731	105	80588		
C x D	2255958	9	250662	6.8455	0.000000
C x D x A	1500543	18	83364	2.2766	0.002450
Error	11534450	315	36617		
B x C x D	175066	9	19452	0.7349	0.676851
2*3*4*1	602885	18	33494	1.2654	0.208768
Error	8337975	315	26470		

Table 3.2: Repeated ANOVA for mean matching RTs was conducted. The same between- and within- participants variables were used as in the ANOVA for error rate.

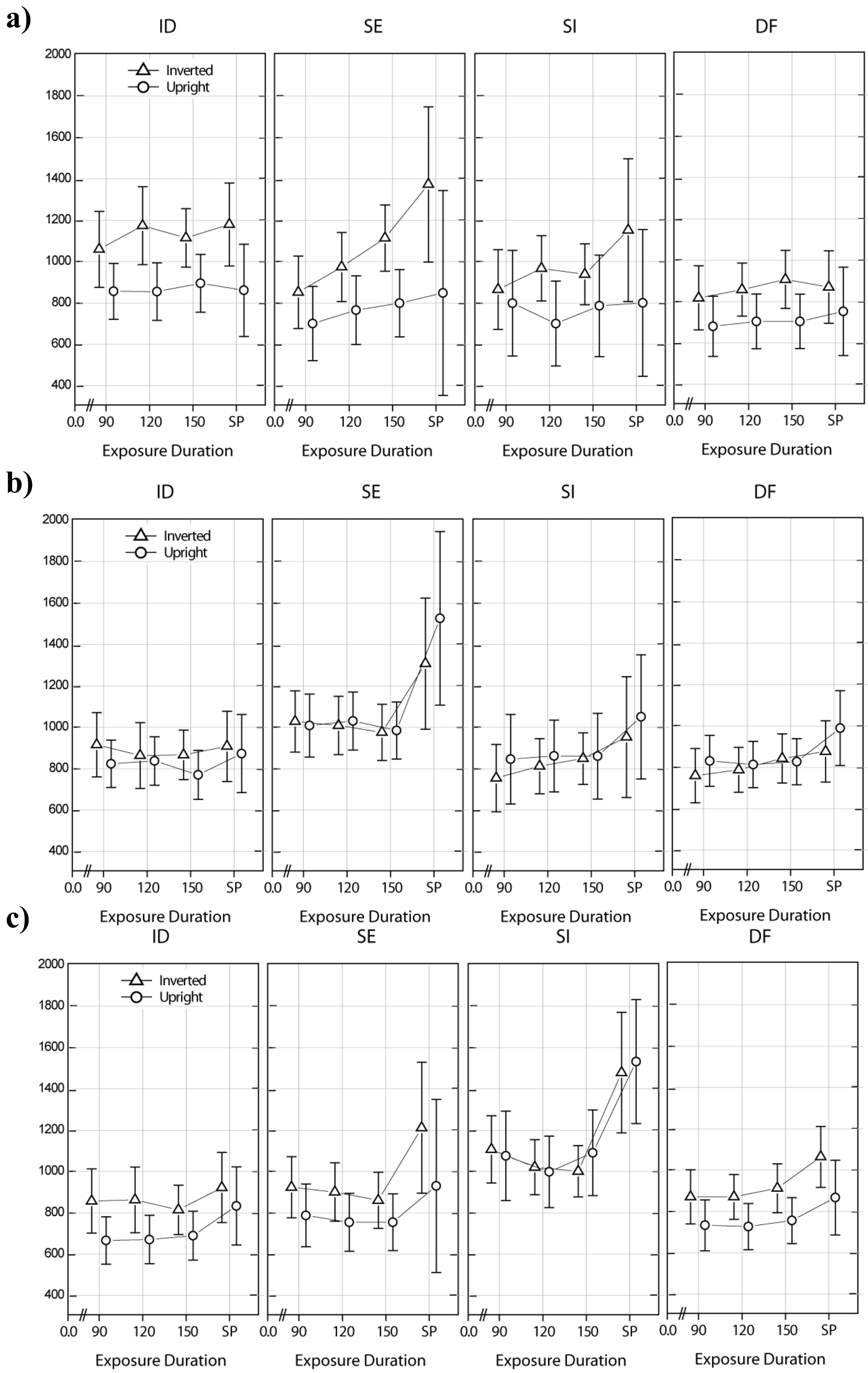

Figure 3.3: Average RTs for a) Task 1 – matching of whole faces, b) Task 2 – matching of external features and c) Task 3 – matching of internal features in upright and inverted orientation. RTs in ms are plotted against the four exposure durations (90 ms, 120 ms, 150 ms, self-paced) and broken up by conditions (ID, SE, SI, DF).

The main effect of exposure duration was also significant with the significantly longer reaction time for self-paced exposure duration than for 90 ms, 120 ms and 150 ms. However, there was a significant two-way interaction between exposure duration and condition as well as a three-way interaction between task, exposure duration and condition. To analyze the three-way interaction, we conducted separate ANOVAs for each of the three tasks: In Task 1 (matching of whole faces), the increase of exposure duration results in longer reaction times, but only for SE-trials ($F(3, 27) = 7.18$, $p < .001$) and SI-trials ($F(3, 27) = 3.08$, $p < .05$), while ID and DF-trials were not influenced by exposure duration (both $p > 1.64$). In Task 2 (matching of external features) the increase of exposure duration results in longer reaction times in the SE, SI and DF-trials (SE: $F(3, 39) = 4.43$, $p < .01$; SI: $F(3, 39) = 3.79$, $p < .05$; DF: $F(3, 39) = 3.67$, $p < .05$), but not in ID-trials ($p = .60$). In Task 3 (matching of internal features), all four conditions (ID, SE, SI, DF) were impaired by exposure duration (all $p < .05$).

3.4.3 Trade-off between reaction times and matching errors

To test a possible speed-accuracy trade-off, correlations between RTs and matching errors were calculated. There were some significant correlations, but all of them positive, showing that speed-accuracy trade-offs did not occur in our data.

3.4.4 Control Experiment

A control experiment was conducted to test the effect of a mask and compare how the lack of a mask influences matching performance in the previous three experiments. Due to long ISI interval and the high cognitive load in our matching task we did not expect any significant impact. In the control experiment the exposure duration was constant at 150 ms, while the test face was resized to be 17% smaller than the target face. After the test face was presented a pattern mask

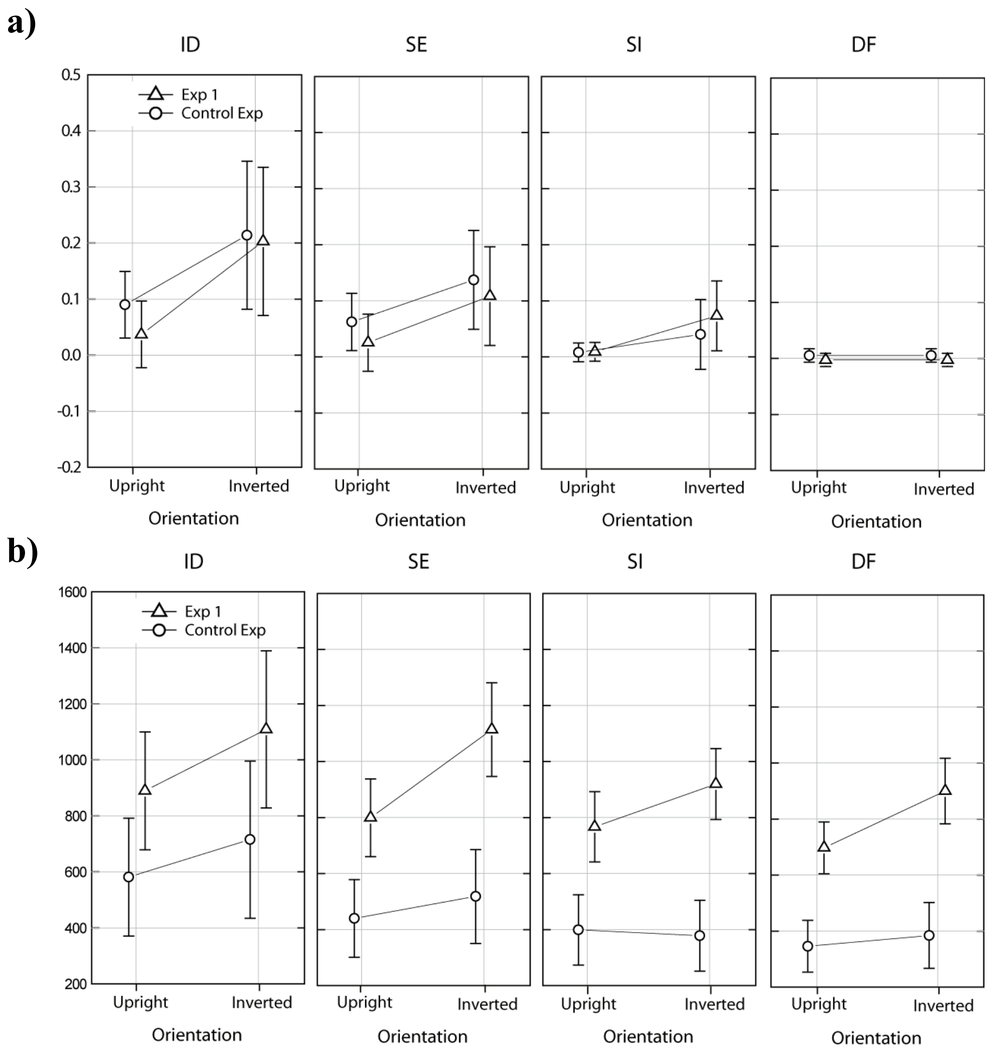

Figure 3.4: Average matching errors (a) and average RTs (b) for Task 1 and a control experiment. Exposure duration in both experiments was 150 ms and the task was to match faces as whole. In the control experiment a pattern mask was presented and the second face resized to be 17% smaller than the target face. No differences in the number of the matching errors was found, although reaction times were faster in control experiment.

was shown for 100 ms. Mean matching errors and mean reaction times from the control experiment were compared with the data for 150 ms exposure duration in Task 1, where whole faces had to be matched. Repeated ANOVA was conducted on mean matching errors and mean reaction times, with Experiment (Task 1 vs. control experiment) as between-participants factor and the Orientation (upright vs.

inverted) and Condition (ID; SE; SI; DF) as within-participants factors. For Error Rate there was no significant difference in matching errors between the two experiments ($F(1, 18) = .51$; $p = .48$), whereas the reaction time was significantly faster in the control experiment where the mask was presented ($F(1, 18) = 24.71$; $p < .001$). We can therefore conclude that the error rate in the control experiment is not caused by longer reaction time. A different computer was used to conduct the control experiment, which may also contribute to the faster reaction time. Nevertheless, the main effects and global trends are the same in both data sets and lead to the conclusion that the lack of the mask did not cause a significant difference in the general performance. Mean matching errors and mean reaction times for both experiments are shown in the Figure 3.4.

3. 5 Discussion

In this study, we investigated the effects of inversion and short exposure duration on the processing of internal and external facial features in the whole facial context. The tasks explored matching of two faces on the basis of their internal features only, of their external features only, or on the basis of both internal and external features (whole face matching).

As the first hypothesis we have tested how orientation impairs matching of whole faces and their facial features. To draw a conclusion about the contribution of the external and internal facial features to the processing of the upright and inverted faces, it seems reasonable to explore it in the context of the whole faces, as we have done it in this study. According to the holistic or configural approach, the inversion effect is expected only for whole faces, but not for facial features (for a review see Valentine, 1988, Schwaninger et al., 2003). However, there are studies revealing an inversion effect when isolated facial features were presented (Rakover & Teucher, 1997; Moscovitch & Moscovitch, 2000; Barton et al, 2001; Nachson &

Shechory, 2002). The data we obtained showed a higher average error rate and longer reaction time for the processing of inverted than for upright faces. Further analysis revealed an inversion effect only for whole face matching (Task 1), while matching of external or internal features remained unimpaired (Task 2 and Task 3). So far, our findings support the hypothesis that inversion impairs holistic or configural processing, while featural processing remains unaffected. Thereby our data seems to contradict earlier studies which imply the same cognitive processes for inverted whole faces as well as for their internal and external features (Rakover & Teucher, 1997; Moscovitch & Moscovitch, 2000; Nachson & Shechory, 2002).

At this point it is important to notice that in our stimuli, the facial proportions within both internal and external features were kept constant in all conditions. In spite of this, the holistic facial pattern changes through a combination of internal and external features in the conditions SE and SI. Moreover, in our inverted condition, both target and test faces were presented inverted (Inverted-Inverted), which leads to weaker inversion effects than if the procedure would have required mental rotation of the features (Upright-Inverted) (Rakover & Teucher, 1997). All of this indicates that facial context, defined as the configurational pattern of all presented features, may have facilitated the extraction of the internal and external features in inverted orientation. That would explain why inversion effect was not obtained for external and internal features in our study, although it was strong (but still weaker than for whole faces) in the studies presenting isolated features. In spite of the fact that methodological procedures could explain some aspects of the data, this assumption has a speculative character and has to be examined in further studies.

A second important aim of our study was to examine effects of exposure duration on the processing of external and internal features. We could demonstrate that within 90 ms exposure duration not only whole faces, but also facial features can be extracted and matched at a level higher than chance. Although reaction times did not differ between the three tasks, the average error was the highest for matching of internal features, while there was no difference in matching of external features and whole faces. These results are in line with the previous findings suggesting more accurate matching on the basis of external rather than internal features, but failed to confirm the faster processing of external compared with internal features (De Haan & Hay, 1986; Young et al., 1985; Nachson et al., 1995).

Interestingly, there was an interaction between exposure duration and orientation suggesting that processing of upright and inverted faces occurs at the different time scales. Namely, the short exposure duration of 90 ms and 120 ms resulted in a higher error rate for inverted, but not for upright stimuli. For upright faces, the maximal performance was achieved at 90 ms exposure duration and longer exposure durations (120 ms, 150 ms, self-placed) did not improve accuracy. For inverted stimuli it seems that a critical exposure duration is met at 150 ms. Further reduction of the exposure duration to 90 ms or 120 ms decreased matching performance significantly. Prolonged exposure duration (self-paced) did not result in a significant increase of matching errors, although there was a tendency in this direction, especially in some conditions. The reaction times in the self-paced exposure duration were significantly longer for both upright and inverted faces, but it did not result in a better performance.

As already mentioned, the composition of a face (ID, SE, SI, DF) is supposed to have a significant impact on face matching (Nachson et al., 1995). Our data confirmed previous results by Nachson et al. (1995) that matching two different faces (condition DF) was the fastest and the most accurate. There was also a significant interaction between exposure duration and condition. Interestingly, our data showed that with longer exposure duration, the reaction time for the SE and SI conditions increased in all three tasks; accuracy, however, did not rise with longer exposure duration. The longer reaction time for SE- and SI- trails, where the facial features have to be isolated from facial context, implicates a predominantly serial and part-based processing. In contrast, when the faces are either completely same or completely different, fast reaction times indicate a more holistic approach

(ID- and DF-trails). The interaction between exposure duration and condition was modified by task, indicating that matching strategies depend on matching criteria (whole face, external or internal features). Generally, the *different* judgments were more accurate and faster than *same* judgments; i.e. in Task 1 (whole face matching), matching of identical faces (ID) was more error prone and slower than the other three conditions (SE, SI, DF). Consequently, in Task 2 (matching of external features) the highest average error score and the longest reaction time were reached with SE-trials and in Task 3 (matching of internal features) with SI-trials.

To summarize, the results in all three tasks provide evidence for very fast face matching. Maximal performance for upright faces was achieved at 90 ms exposure duration, while longer exposure durations, albeit leading to longer reaction times, did not reduce the error rate. For inverted faces, performance continually decreased with exposure durations below 150 ms. Our results are in line with previous findings suggesting that matching of unfamiliar faces is more accurate on the basis of external rather than internal features. Contrary to other studies, we could not confirm faster processing of external than internal features. However, the data implies a part-based and time consuming process for both internal and external features if they have to be matched within facial context. Finally, this study demonstrated that manipulating exposure duration reveals some evidence about the time course in face matching.

4. Does Matching of Internal and External Facial Features Depend on Viewpoint?

4.1 Summary

To test whether internal (eyes, eyebrow, nose and mouth) and external (hair, ears, head and face outline) facial features are affected by viewpoint, orientation and exposure duration in the same way, we conducted two sequential matching tasks. Matching criteria were based on either external or internal facial features, which were manipulated in a whole facial context. The results of the two conducted experiments have revealed that processing of internal facial features is highly sensitive to view, orientation, and time. External facial features are found to be processed much more robustly. There were no orientation or viewpoint effects for matching of external features. According to our results, external features might be processed more like objects, while internal features contain face-specific information.

4.2 Introduction

The human cognitive system deals with a large number of objects and faces, recognizing them readily and effortlessly. This phenomenon is even more intriguing if one takes into account the many views which can be produced of each single item. The topic has been investigated extensively, but with variable and inconclusive results. For this reason, the discussion about view-dependent or view-independent object processing has shifted to a new focus. Multiple routes to object recognition, which can be situated on a continuum ranging from viewpoint dependency to complete viewpoint independency, have been adopted in recent years in vision science. Moreover, it has been revealed that stimuli and the task have a substantial impact on view-dependent or view-independent processing (Vanrie, Béatse, Wagemans, Sunaert, & Van Hecke, 2002; Biederman & Bar, 1999; Edelman, 1995; Hayward & Tarr, 1997; Hayward & Williams, 2000). The mental rotation process is an extreme case of viewpoint dependency; response times in a same/different judgment increase linearly with increasing angular difference between two comparison objects (Shepard & Metzler, 1971). In contrast, using locally diagnostic, view-invariant features to determine the identity of an object enables viewpoint-independent processing (e.g., Eley, 1982).

In both face and object recognition, a ¾ view is supposed to be a canonical view in which object and faces can be better represented and more easily recognized (Palmer et al., 1981; Verfaillie & Boutsen, 1995; Bruce et al., 1987). The advantage of a ¾ view in face processing, it is considered, is that it provides information about both frontal and profile views (e.g. Baddeley & Woodhead, 1983; Fagan, 1979). However, this advantage has been found almost exclusively in face recognition tasks, where subjects are shown target faces during the learning phases and are then required to identify those faces among distracters in test phases. In contrast, studies with matching tasks have identified angular difference as a predictor for performance in matching two faces from different viewpoints. It has been shown that two arbitrary views with smaller angular differences can be generalized better than two views with a larger one (Lee, Matsumiya, & Wilson, 2006, Liu & Chaudhuri, 2002). It has also been found that the viewing angle affects the matching of unfamiliar faces much more than that of familiar faces (Bruck, Cavanagh, & Ceci, 1991; Davies & Milne, 1982; Hill & Bruce, 1996).

How processing of facial features depends on viewpoint has not, however, been the explicit topic of previous studies, although there is evidence for a substantially different contribution of internal and external features to face processing (Elis et al., 1979; Hines et al., 1987; De Haan & Hay, 1986; Young et al., 1985; Nachson et al., 1995; Bruce et al., 1999). Internal features (mouth, nose, eyes, eyebrows) are found to be most important for matching of familiar faces (Elis et al., 1979; Hines et al., 1987), while external features (hair, ears, head and face outline) are related predominantly to the matching of unfamiliar faces (De Haan & Hay, 1986; Young et al., 1985; Nachson et al., 1995; Bruce et al., 1999; Veres-Injac & Schwaninger, in press). Furthermore, it has been proposed that processing of internal and external features is mediated by different parts of the neural system. The face recognition system performs mental representation of known faces based largely on internal features (Moscovitch & Moscovitch, 2000; Bonner, Burton, & Bruce, 2003), while external features are processed predominantly by the object recognition system (Moscovitch & Moscovitch, 2000).

The face recognition system consists of three bilateral regions in the occipitotemporal visual extrastriate cortex and includes inferior occipital gyri, the lateral fusiform gyrus and the superior temporal suclus (for a review see Haxby, Hoffman, & Gobbini, 2000). The object recognition system includes regions which are located near the face-selective fusiform region in the parahippocampal, fusiform and inferior temporal gyri. These regions respond more to object categories (e.g. houses, chairs, tools) than to faces (Haxby et al., 1999; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Aguirre, Zarahn, & D'Esposito, 1998). Moscovitch and colleagues (1997) in a study with CK, a person with object agnosia, found that intact face perception mechanisms by themselves cannot process inverted faces effectively and that human neural systems for face and object perception interact to accomplish inverted face perception. In a second study with CK, Moscovitch and Moscovitch (2000) found that CK recognized faces normally from their internal features only, but recognition from purely external features was much worse in comparison with the subject control group, who performed equally well on both internal and external features. This may be expected only if external features are processed more effectively by the object recognition system and internal features by the face recognition system, since the CK`s face recognition system is intact but his object recognition system is damaged. The authors reason that the face recognition system operates primarily on holistic properties of faces and is predominantly sensitive to the configuration of upright internal features. The object recognition system operates primarily on facial components or features and uses part-based information to integrate individual features into whole facial representation (Moscovitch & Moscovitch, 2000).

In this study we test whether matching of external and internal facial features is view-dependent. Mental rotation is required in order to complete a serial matching of two faces from different views. However, if only matching of either internal or external features has to be complete, isolated facial features can be used as matching criteria. According to current knowledge about object recognition, isolated features can be used as criteria for object recognition independent of viewpoint. If external features are processed part-based and predominantly by the object recognition system (Moscovitch & Moscovitch, 2000), we may expect no viewpoint or inversion effect when faces are to be matched only by external features. However, internal features are expected to be impaired by both viewpoint and orientation, since internal features contain configural information and are processed by the face recognition system. To test this hypothesis we conducted two experiments with a serial matching task in which whole faces or facial features from frontal and ¾ views had to be matched.

The effects of orientation and exposure duration on matching performance are also important issues which this study addresses. Our previous findings suggested we should observe fast and accurate matching of upright faces in frontal view (Veres-Injac & Schwaninger, in press). Although exposure duration was reduced from self-paced to 90 ms, we did not find a difference in matching errors between four different durations (90 ms, 120 ms, 150 ms and self-paced). However, matching errors for whole inverted faces increased when exposure duration was set below 150 ms. There is also evidence from neuropsychological studies showing that reduction of ERP response for identical inverted faces is delayed by 30 ms in comparison with identical upright faces (Jacques et al., 2007). Both studies seem to lead to the conclusion that more time is needed to process inverted faces. There is, however, no agreement on whether inversion equally impairs matching of whole faces and facial features (Valentine, 1988; Schwaninger et al.; Rakover & Teucher, 1997; Nachson & Shechory, 2002). Our previous results tend to support the hypothesis that inversion impairs holistic or configural processing, while featural processing remains unaffected (for a review see Valentine, 1988; Schwaninger et al., 2003). Before this can be definitely claimed it is important to notice that in the previous study *both* target and test faces were presented inverted in the inverted condition. Strictly speaking, this means that faces in inverted orientation might have been compared as unparsed perceptual wholes; mental rotation of facial features was not required to solve the task. In this study we employed a sequential matching task, where the first face was always in upright orientation, while the second (test) face could be present either upright or inverted. Our aim was to test whether matching of facial features and whole faces is similarly impaired by orientation when solving the task requires mental rotation. Moreover, as mentioned above, matching of whole faces in upright and inverted orientation is time sensitive (Jacques et al., 2007; Veres-Injac

& Schwaninger, in press). In line with those findings, we propose that matching from different viewpoints could also depend on exposure duration. However, if external and internal facial features are processed independently from viewpoint, then the duration for facial features might not differ in frontal and ¾ views.

To test our hypotheses, we measured behavioral performance in two different matching tasks: matching of only external (Experiment 1) or only internal (Experiment 2) features. The IDE condition was present in both experiments. This meant that the decision about face identity could be made on the basis of the whole face, since there was either full identity (in both internal and external features) or full difference between target and test faces. However, the SISE condition, in which the decision about face identity could be made only onto basis of either internal or external features matching, was also present. The reaction times and matching errors were collected for faces matched in frontal and ¾ views. Stimuli were shown in upright and inverted orientation. A sequential matching task was used, with constant exposure duration for the target faces (1500 ms), and two different exposure durations for the test faces (60 ms, 150 ms).

4.3 Materials and Procedure

4.3.1 Participants

A total of 60 students of psychology participated in the two experiments reported in this study, 30 in Experiment 1 (17 females), and 30 in Experiment 2 (19 females). All participants had normal or corrected to normal vision.

4.3.2 Stimuli

Full-color frontal and ¾ view photographs of twelve male faces were captured in a photo studio under controlled lighting conditions. None of the faces were wearing glasses or jewelry or had a beard. The haircut of all faces was kept comparably similar, with a short hairstyle. Each of the four target faces was paired with the appropriate 12 test faces in frontal view and the same 12 faces in ¾ view. The total of possible pairs were 96. The test faces were assigned to four conditions. In the identical (ID) condition, the same face was used as test and target face. In the same external (SE) condition, both target and test faces shared the same external features, while internal features differed. In the same internal (SI) condition, only the internal features of the target and test faces were the same. In the different (DF) condition, target and test faces depicted two different individuals with no common features at all. For the purposes of statistical analysis, the wholly identical (ID) and wholly different faces (DF) are combined in one condition, (IDE), since the decision about face identity can be made from the whole face. Faces sharing only the same external (SE) or only the same internal features (SI) were combined in the SISE condition, since the decision about face identity can be made only through part-based matching. The test faces for SE and SI conditions were prepared combining external and internal features of the test and target faces in Adobe Photoshop 9. The internal features (nose, mouth, eyes, eyebrows) were cut out with comparable tracing lines and placed on the second (template) face, based on the position of the internal features (Veres-Injac & Schwaninger, in press). Figure 4.1 shows a target face and four test faces in two different views.

All stimuli were 300 x 400 pixels in size. Images were presented on a 17’’ color CRT monitor. The presentation position of the first face was always shifted by -20 pixels from the center, while the second face was always shifted in the other direction by +20 pixels. Screen resolution was set to 1024 x 768 pixels with a refresh rate of 60 Hz. The viewing distance was approximately 60 cm.

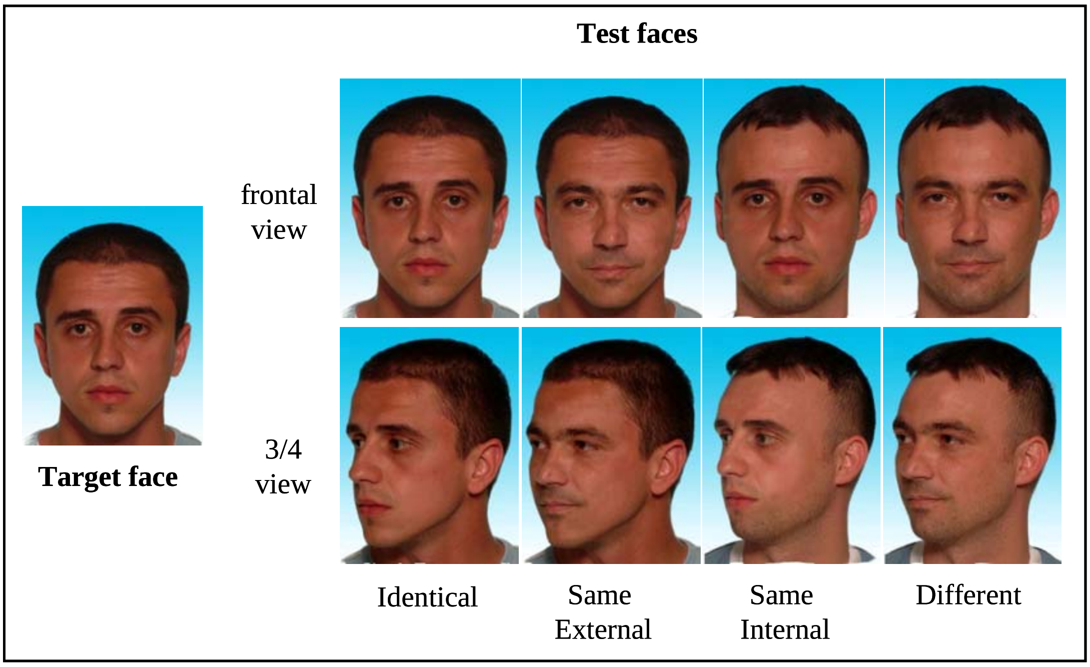

Figure 4.1: Four different types of stimuli, which were presented in Experiment 1 and 2. Identical and Different faces can be processed as unimpaired wholes and together are represented in the condition IDE. Contrary, Same External and Same Internal faces share only partially the same features. Together they form condition SISE.

4.3.3 Experimental design

Each trial started with a target face in frontal view that was displayed for 1500 ms. After a blank screen of 1000 ms, a test face appeared. A total of 384 trails were presented in each experiment. In half of these, test faces were shown in frontal view and, in the other half, in ¾ view. Target faces were always presented in frontal view and upright orientation, whereas the orientation, view and exposure duration of test faces varied. Training preceded the main experiments to make sure that the subjects understood the task. Training consisted of 8 trails in both views, with additional facial stimuli which were presented only in training sessions but not in the main experiment. Participants were asked to respond whether the target and the test faces were the same or different by pressing the left or right mouse button with their preferred hand as quickly and accurately as possible. The assignment of answers (*same*/*different*) to the left or right mouse button was counterbalanced across participants. The definition of *same* was varied between Experiments 1-2. In Experiment 1 faces had to be matched on the basis of external features, while in Experiment 2 the matching criteria were internal features. In both experiments, however, the IDE condition, where faces as whole can be matched - i.e. where internal and external features are congruent and either identical or different with the target face - operated. In the SISE condition, faces only share the same internal features, whereas external are different or the same external features, with different internal features (see Figure 4.1.). The correct decision about face identity in the SISE condition can be made correctly only if part-based and depending on the definition of *same*. Test faces were presented in two different exposure durations, 60 and 150 ms. The exposure duration of test faces was randomized across trials.

4.4 Results

4.4.1 Mean correct response times

Reaction times in Experiment 1 and 2 were analyzed by 4-way repeated measures of ANOVA. The Task (matching of internal or external facial features) was a between-subject variable, while Condition (IDE – matching of whole faces; SISE – part based matching, i.e. matching of isolated features), Orientation (Upright; Inverted), View (Frontal, ¾ view) and Exposure duration (60 ms, 150 ms) were within-subject variables. Significant main effects and interactions are showed in Table 4.1.

Reaction times for upright-presented faces were faster than for inverted-presented faces. Also, whole-face matching (IDE condition) was completed significantly faster than part-based matching (SISE condition). Since there was significant interaction between Orientation and Condition, additional pairwise comparisons were conducted, which have revealed longer reaction times for inverted faces if they are matched as wholes (IDE), $F(1,58) = 19.54$, $p < .001$. There was no significant inversion effect on reaction times for matching of facial features (SISE condition), $F(1,58) = 0.37$, $p = .54$.

Source	**SS**	**df**	**MS**	**F**	**p**	**Partial Eta Squared**
ORIENTATION	86717.74	1	86717.74	5.477	0.05	0.086
Error	918304	58	15832.83			
CONDITION	1293310	1	1293310	49.723	0.001	0.462
Error	1508589	58	26010.15			
VIEWPOINT x TASK	118405.5	1	118405.5	4.73	0.05	0.075
Error	1451921	58	25033.13			
ORIENTATION x VIEWOINT	60424.91	1	60424.91	4.116	0.05	0.066
Error	851458.6	58	14680.32			
ORIENTATION x CONDITION	156440.5	1	156440.5	15.945	0.001	0.216
Error	569061.4	58	9811.403			
VIEWPOINT x CONDITION	162715	1	162715	9.998	0.01	0.147
Error	943916.7	58	16274.43			

Table 4.1: Significat main effects and interactions for mean RTs analysis. The table shows source of variation, sum of squares (SS), degree of freedom (df), mean of squares (MS), *F*-ratio, significance level (p), and the ratio of explained of total variation.

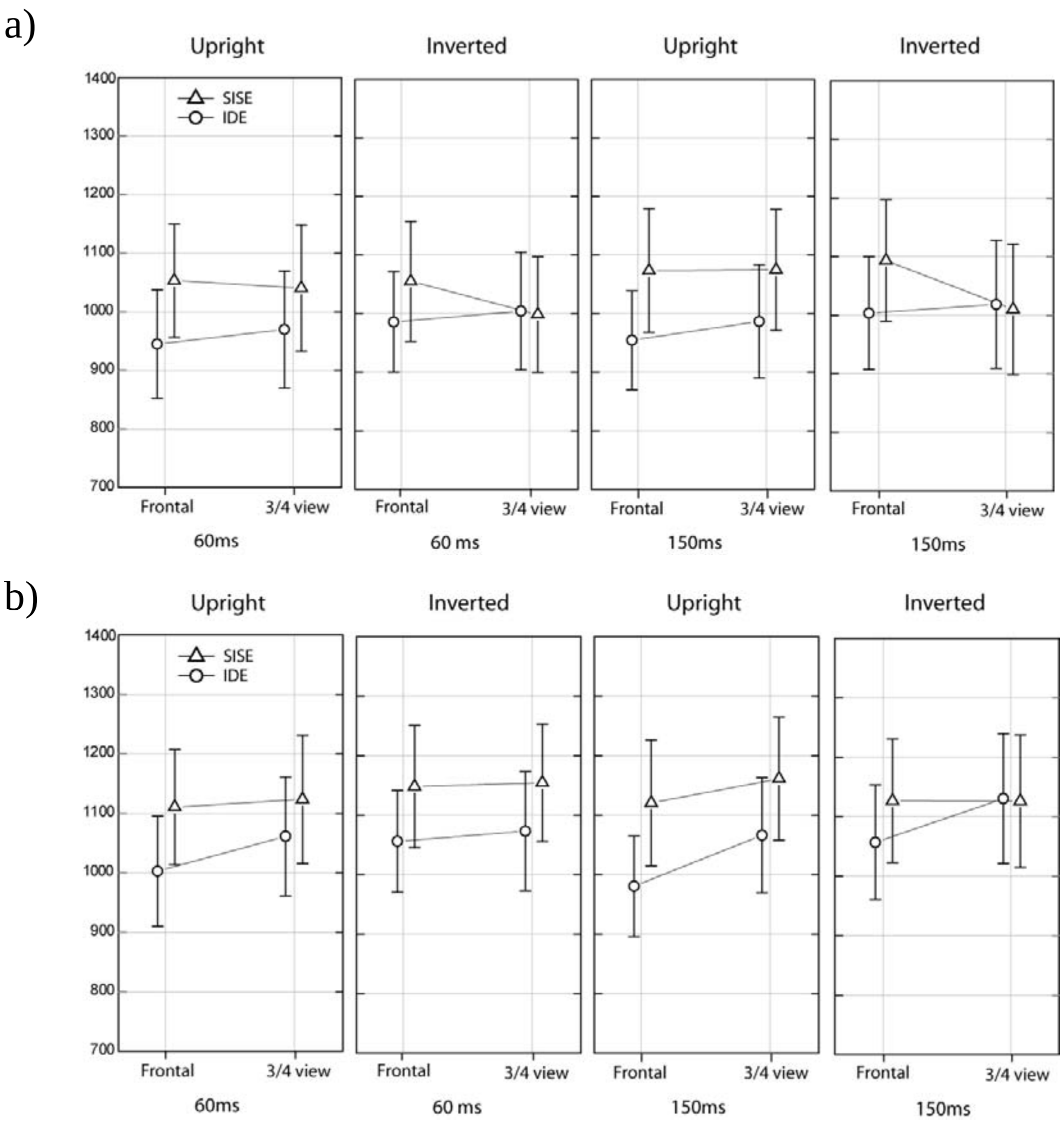

Figure 4.2: Average reaction times in milliseconds in a) Experiment 1 – matching of external features and b) Experiment 2 – matching of internal features, in upright and inverted orientation. Average reaction times for IDE and SISE conditions are plotted against the two viewpoints (frontal and ¾ view) and broken up by exposure duration (60/150 ms).

In Experiment 1, where matching of external features was required, reaction time for faces in frontal and ¾ view was equally fast, $F(1,58) = 0.27$, $p = .60$. Viewpoint effect was significant in Experiment 2, where matching criteria were internal features. Matching of internal facial features from ¾ view compared with frontal view lasted significantly longer, $F(1,58) = 6.50$, $p < .05$. Moreover, reaction times for part-based matching in the SISE condition were not affected by viewpoint, $F(1,58) = .60$, $p = .44$, whereas reaction times for IDE condition were significantly longer if faces were presented in ¾ view, $F(1,58) = 12.90$, $p < .001$. Mean reaction times for Experiment 1 and 2 are shown in Figure 4.2.

Source	**SS**	**df**	**MS**	**F**	**p**	**Partial Eta Squared**
EXP. DURATION (A)	0.05658	1	0.05658	10.2744	0.01	.150
Error	0.31940	58	0.00551			
ORIENTATION (B)	1.10908	1	1.10908	88.1739	0.001	.603
B x TASK	0.25470	1	0.25470	20.2489	0.001	.259
Error	0.72954	58	0.01258			
VIEWPOINT (C)	1.10840	1	1.10840	83.0694	0.001	.589
C x TASK	0.27903	1	0.27903	20.9123	0.001	.265
Error	0.77390	58	0.01334			
CONDITION (D)	5.92024	1	5.92024	127.6372	0.001	.688
Error	2.69023	58	0.04638			
A x B	0.02471	1	0.02471	4.1692	0.05	.067
Error	0.34369	58	0.00593			
B x D x TASK	0.17726	1	0.17726	15.0511	0.001	.206
Error	0.68308	58	0.01178			
C x D x TASK	0.20402	1	0.20402	16.2557	0.001	.219
Error	0.72793	58	0.01255			
B x C x D	0.02972	1	0.02972	4.6073	0.05	.074
Error	0.37413	58	0.00645			

Table 4.2: Significant main effects and interactions for mean matching errors analysis. The table shows source of variation, sum of squares (SS), degree of freedom (df), mean of squares (MS), F-ratio, significance level (p), and the ratio of explained of total variation.

4.4.2 Mean matching errors

The statistical procedure and between- and within-subject variables was the same as for the reaction times analysis. Table 4.2 shows statistically significant main effects and interactions for matching errors.

Further, pairwise comparisons were conducted in order to analyze interactions which were of interest in this study: interaction between Viewpoint, Condition and Task, interaction between Orientation, Condition and Task, and interaction between Exposure Duration and Orientation.

We found a significant effect of Viewpoint for whole-face matching (IDE) in both tasks, where the mean number of matching errors was larger for ¾ view than for frontal view (both $p < .001$). However, a similar effect for part-based face matching (SISE) was only found in Experiment 2, where internal features had to be matched, $F(1,58) = 51.67$, $p < .001$. If part-based matching of external features was required (SISE condition in Experiment 1), there was no difference in matching errors in frontal and ¾ view, $F(1,58) = 0.22$, $p = .63$.

Further analyses referring to the orientation effect yielded a similar pattern of results. A significant increase in mean matching errors in both experiments was found for whole face matching (IDE), both $p < .001$. Also, orientation significantly impaired part-based matching of internal facial features (SISE) in Experiment 2, where the mean matching errors increased significantly for inverted faces, $F(1, 58) = 62.73$, $p < .001$. Again, part-based matching of external features in Experiment 1 was not impaired by orientation, and there was no increase in mean matching errors in inverted orientation, $F(1,58) = .004$, $p = .83$. An interaction between Exposure Duration and Orientation also reached significance. Pairwise comparisons revealed that a decrease of exposure duration from 150 ms to 60 ms increased the number of matching errors in both experiments, but only if faces were presented inverted (both $p < .05$). If presented upright, face matching remained unimpaired by Exposure Duration. No further significant interactions between Exposure Duration and other factors were found. Mean matching errors in Experiment 1 and 2 are shown in Figure 4.3.

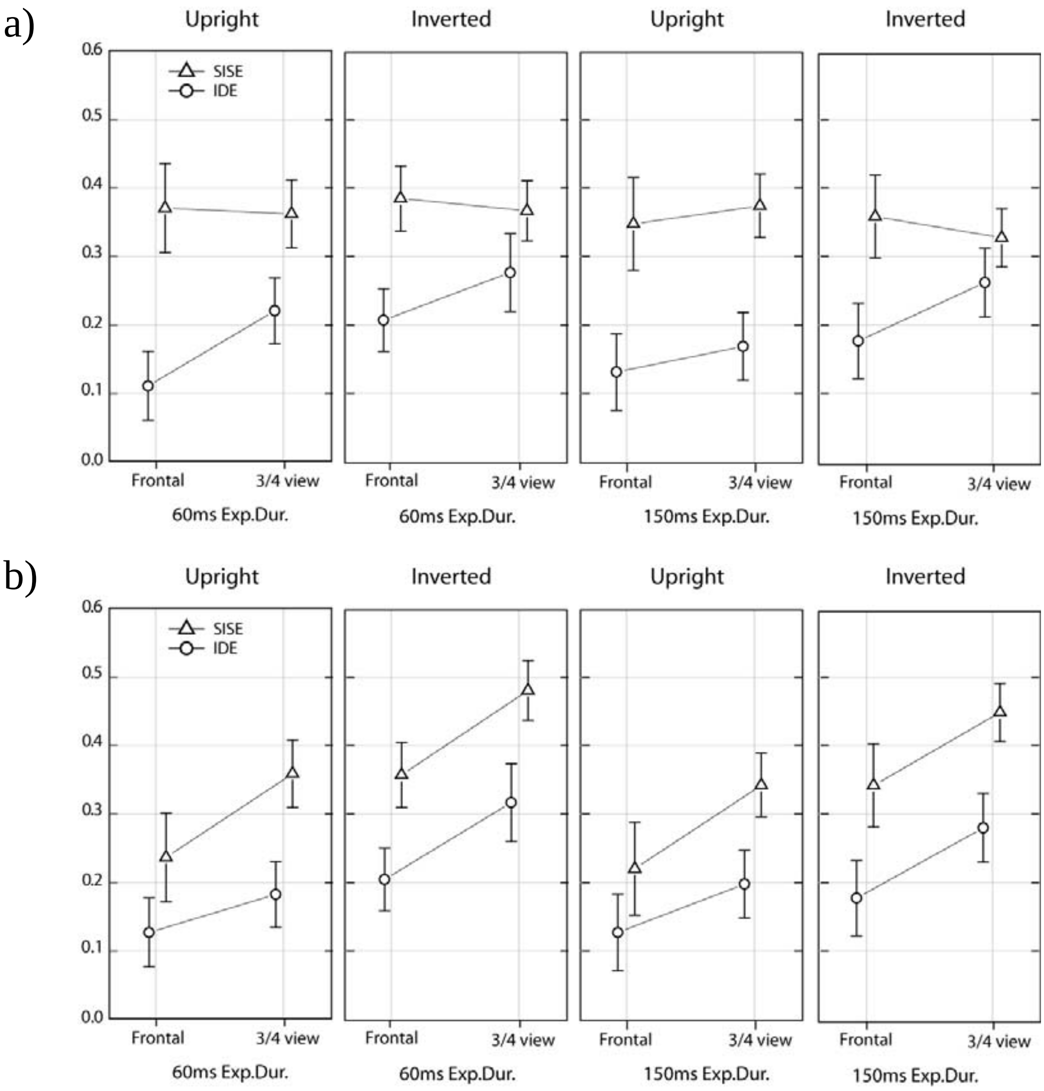

Figure 4.3: Percent of average matching errors in: a) Experiment 1 - matching of external features and b) Experiment 2 - matching of internal features, in upright and inverted orientation. Average error rates for IDE and SISE conditions are plotted against the two viewpoints (frontal and 3/view) and broken up by exposure duration (60/150 ms).

4.5 Discussion

In this study we addressed three questions relevant to the processing of faces on the basis of internal and external facial features. The first question concerns whether matching of facial features is view dependent. In accordance with the data observed in object recognition studies one might suppose that facial features are processed independently of viewpoint. This appears to be particularly justifiable for external features, which are supposed to be processed predominantly by the object-recognition system (Moscovitch & Moscovitch, 2000). To test this we compared matching of whole faces, and of external and internal facial features form frontal and ¾ views in two experiments. In Experiment 1 only matching of external facial features (hair, ears, head and face outline) was required. In Experiment 2 faces were matched only on the basis of internal features (eyes, eyebrows, nose and mouth). In both experiments there were two conditions. The first was IDE, where target and test faces shared either all or no common characteristics, and therefore could be matched as unimpaired wholes. In the SISE condition target and test faces shared either same internal or same external features, and could be correctly matched only part-based, and depending on the matching criteria. Therefore we reasoned that performance in the SISE condition may reflect processing of external and internal features in Experiment 1 and 2 respectively. Further, we expected that processing in the IDE condition will rely on all available information (i.e. whole facial context).

Generally, we were not interested in whether faces are matched faster and more accurately in the frontal or ¾ view, but rather in revealing how angular change from one to another view affects matching performance for external and internal facial features. Mental rotation is required to match faces from two different views, as well as for matching two faces in different orientations. Accordingly, matching faces in the same view and in upright orientation does not require such a mental rotation and is expected to be completed more easily.

However, if external and internal features are processed differently, with the first one relying more on the object-recognition system and the second on the face-recognition system, we may expect that effects like viewpoint and orientation impair performance differently –depending on which kind of features has to be matched.

The results of this study support a claim proposed by Moscovitch and Moscovitch (2000) that internal and external facial features might be processed differentially. A strong viewpoint effect was found for matching of internal, but not of external, features. Whereas mental rotation from one to another view for internal features causes an increase in reaction times and matching errors, no such effect was found for external features. Indeed, reaction times and matching errors for external features remained astonishingly stable in both frontal and ¾ views. Matching of whole faces was found to be viewpoint dependent. However, whole faces are generally processed much faster and more accurately than isolated internal and external features, suggesting that whole-face processing benefits from information from both kinds of features.

The second question was related to the effect of orientation for external and internal facial features. Previous studies showed a different pattern of results, with two opposite claims of orientation independency and orientation dependency in the processing of facial features (Valentine, 1988; Schwaninger et al., 2003; Rakover & Teucher, 1997; Nachson & Shechory, 2002). Our previous study (Veres-Injac & Schwaninger, in press) corroborates a rather holistic hypothesis, suggesting that facial features are not impaired by inversion. Due to the methodologically based preference for holistic processing in inverted orientation, we were restricted in the generalization of results in that previous study. Our approach, however, has a very important advantage compared to those used in other studies; it enables us to investigate processing of facial features in the context of whole faces, adding ecological validity to our results. In this study we aimed to test whether facial features became orientation-sensitive if mental rotation of features was required. To study this, two faces, the first upright followed by an inverted face, had to be matched. We have found an inversion effect for whole faces and internal facial features in this task. For whole faces, reaction times, as well as matching errors were affected by inversion, and increased significantly compared with upright faces. Matching of inverted internal facial features did not result in longer reaction times, but in an increase of matching errors. Matching of external facial features remained unimpaired by orientation; there was no increase in reaction times or in matching errors. The lack of an orientation effect for external features and its strong impairment on internal features might be interpreted in favor of the hypothesis that external and internal features are processed by two different sensory systems, one of which is more object-specific and the other more face-specific.

The third question addressed in this study is a relatively new and insufficiently investigated aspect of duration in face processing. Our point of origin was some findings suggesting that a longer time frame is needed for matching inverted than for matching upright faces (Jacques et al., 2007; Veres-Injac & Schwaninger, in press). We supposed that the duration of face processing depends on the complexity of the cognitive steps that have to be conducted in order to complete the task (Sternberg, 1969, 1998a, 1998b). In other words, the performance is expected to be impaired by shortening exposure duration. How large the decrease in performance will be depends on task complexity. To test this hypothesis, we used two exposure durations of test faces, 60 ms and 150 ms. These critical durations were chosen to challenge performance in a larger spread. In fact, the previous study showed no difference in performance between 90 ms and 150 ms for upright faces (Veres-Injac & Schwaninger, in press), indicating that shorter exposure durations would have to be employed. However, exposure duration below 150 ms for inverted faces caused an increase in error rate; that is why we chose this exposure duration as an upper time limit. The data we observed in this study enhance our previous results, and demonstrate that upright faces can be successfully matched as fast as within 60 ms of exposure duration. Decrease in exposure duration from 150 ms to 60 ms only resulted in a significant increase in matching errors if faces were presented inverted. However, the effect of exposure duration did not interact with task or viewpoint, suggesting that internal and external feature matching was similarly impaired by exposure duration in both frontal and ¾ view.

In summary, the results of the three experiments of this study provide evidence for a highly view-sensitive and orientation-sensitive processing of internal facial features. In contrast, external facial features have been found to be processed much more robustly. The lack of an orientation and a viewpoint effect in matching of external features suggests that internal and external facial features might be treated differently. According to our results, it seems that external features may be processed more like an object, while internal features bear face-specific information. This assumption directly corroborates the hypothesis postulated by Moscovitch and Moscovitch (2000) that the object recognition and face recognition systems are both involved in face processing. Moreover, our study has shown that internal and external facial feature matching was affected by a decrease of exposure duration below 150 ms. Short exposure duration (60 ms) triggered an increased number of matching errors in both tasks for inverted-presented faces. However, further research is needed to improve our knowledge about the timing of facial features processing.

Generally, the best matching performance was achieved in the whole-faces condition, where both internal and external features were congruent. A plausible explanation for this whole-to-part advantage is contributed by other whole-part experiments (Tanaka & Farah, 1993; Tanaka & Sengco, 1997) and experiments with composite faces (Hole, 1994; Hole, George, & Dunsmore, 1999; Young et al., 1987). The results of these studies suggest that spatial interactions within internal features, and among internal features and external features influence perception of facial features, or halves of faces. Our study has shown that the same spatial or holistic mechanisms influence matching of internal and external facial features when they are manipulated within the whole facial context.

5. Timing of Internal and External Facial Features

5.1 Summary

Timing of face processing was studied by measuring proportion correct functions of exposure duration in a face matching task with upright and inverted faces. Subjects were instructed to attend either internal features (task A) or external features (task B) and matched two consecutive test faces, which were either completely identical or completely different, or they coincided just in the features to be attended. Matching of external features was fast, reaching a 75% correctness level within the first 85 ms, independent of face inversion or the degree of feature congruence. In contrast, matching of internal features at the same level of accuracy took more double this time, and was particularly slow when the test faces coincided only in the features to be attended. We obtained a pronounced face inversion effect at brief timings, which vanished after about 120 ms for matching of external features, but remained at a high and constant level for matching of internal features beyond durations of 200 ms. Our results indicate that external features are processed fast and precede internal features in the stream of processing. The findings of poor performance with just internal features for short exposure durations below 120 ms but high performance and strong sensitivity to face inversion for exposure durations beyond 200 ms suggests that configural information starts to become effective at the earliest after about 120 ms, and is well developed when the first 200 ms have passed.

5.2 Introduction

Humans can easily distinguish faces, recognize them as known or unknown, as happy or sad, as friend or enemy. Even more amazing than our ability to ”read” peoples’ faces and their emotional or motivational meanings is that we are able to do this immediately and effortlessly. The sources of information that underlie this ability and the ways they interact are in the focus of current debate. Consensus is reached that there are at least two classes of information, featural and configural, each providing a relevant contribution to what we refer to when we are aware of a human face. Facial features, understood as isolated and more or less independent facial parts (i.e. hair, eyebrow, eyes, nose, mouth, cheeks and chin), can be easily distinguished, and seem to be weighted differently, where hair, face outline, eyes and mouth are usually considered as determining perception and memory of faces (McKelvie, 1976; Shepherd et al., 1981; Fraser et al., 1990). Internal features (eyes, eyebrows, nose, mouth) are particularly focused, being fixated first and attracting most of the gaze time in scan path measurements with human faces, while external features (face outline, hairs and ears) are seemingly not in the focus of active viewing (Williams & Henderson, 2007; Henderson, Falk, Minut, Dyer, & Mahadevan, 2001).

Another kind of information contained in faces is the spatial arrangement of facial features. This ”configural” information is typical of faces as a homogenous class of objects sharing the same kind of feature information in prototypical arrangement (i.e. eyes are always above nose, nose is above mouth etc.). The small variations in the spatial distances between the facial features are unique for each face, and are often considered as more informative than the characteristics of the facial features themselves (Diamond & Carey, 1986; Rhodes, 1988).

Although there is a general agreement that both kinds of information are relevant and processed by the visual system, current approaches differ with respect to the involved processing steps and their temporal order in the stream of face perception. Holistical approaches propose a simultaneous and highly dependent processing of both featural and configural facial properties (Ellis, 1975; Tanaka & Farah, 1993; Farah et al, 1995), maintaining that all kinds of facial features are perceptually available at one moment in time. Making isolated features explicit is seen as an additional process requiring extra time and resources. Other approaches, stressing either configural or featural information, suggest more or less independent processing stages with different weighting of the two sources of information (Walker-Smith, 1978; Tversky & Krantz, 1969; Macho & Leder, 1998; Bartlett & Searcy, 1993; Diamond & Carey, 1986; Rhodes et al., 1993; Searcy & Bartlett, 1996; Carbon & Leder, 2005). It is claimed that different stimulus aspects are processed at different moments in time, accumulating the final perceptual result. In feature based approaches facial feature processing is conceived as being fast, preceding configural information which binds the parts into perceptual wholes (Bachmann, 1991; Macho & Leder, 1998; Carbon & Leder, 2005). This leans on traditional theories of vision, which assume a part or feature based analysis stage in distal areas of the visual system as coming first, with immediate and effortless feature registration (Marr, 1982; Treisman & Gelade, 1980; Biederman, 1987).

In an interesting face recognition experiment Carbon & Leder (2005) used familiar faces and “thatcherised” versions of these faces, with eyes and mouth 180^{o} rotated within the face. The latter kind of stimuli appears deranged, evoking the impression of grotesqueness. However, when turned upside down, they appear as rather normal faces, and the distorted configural relationships are noticed only by detailed inspection. This has prompted the interpretation that processing of configural information is disrupted in inverted faces, letting the observer rely on just featural information in his judgements (Leder, Candrian, Huber, & Bruce, 2001; Bartlett & Searcy, 1993; Bartlett et al., 2003; Searcy & Bartlett, 1996; Rhodes et al., 1993; Leder & Bruce, 1998, 2000; Rossion & Boremanse, 2008). Now, Carbon & Leder (2005) compared recognition performance of normal and thatcherised faces, both turned upside down (inverted). They found that at brief presentation times (26 ms, masked) subjects performed better with inverted thatcherised faces than with inverted normal faces. At more relaxed timings (200 ms, masked) the effect reversed, and subjects were better with inverted normal faces. Performance with inverted thatcherised faces was constant across presentation times. Just performance with inverted normal faces improved at the relaxed timing. As an interpretation the authors proposed that different kinds of information are available at different moments in time. At brief timings configural information has not yet established, and the observer relies just on featural information. Since eyes and mouth are in their normal orientation in thatcherised inverted faces, a match of these is fast and accurate. Later, configural information adds, however, leading to improvement of just the intact facial stimuli (Carbon & Leder, 2005).

Although tempting, the conclusion that featural information precedes configural information in the task at hand hinges on several assumptions. First, it is assumed that a face inversion effect (FIE) exists for isolated features at brief timings, since the advantage of thatcherised faces at brief timings is explained with a FIE for eyes and mouth. Indeed, there is some evidence for a (small) FIE with isolated facial features (Rakover & Teucher, 1997; Barton, et al., 2001; Malcolm, Leung, & Barton, 2005). However, the timing prerequisites of this effect are not yet clear. Strong FIE modulations are obtained by contrasting internal and external facial features. Perception of sets of external features is generally less affected by inversion than perception of sets of internal features, indicating that the latter convey a substantial amount of configural information when presented upright, i.e. in the usual spatial arrangement congruent with everyday visual experience (Moscovitch & Moscovitch, 2000; Phillips, 1979; Nachson & Shechory, 2002; Veres-Injac & Schwaninger, in press). As shown by a control experiment, the advantage of thatcherised faces at brief timings was only present with whole faces, but vanished when sets of internal features were used (Carbon & Leder, 2005).

Within the authors' reasoning this result is unexpected and raises problems for the explanation of a featural temporal precedence over configural information.

A possible approach to the question at hand is to determine the timings of internal and external features in upright and inverted presentation, and in their natural facial context. Present findings suggest that internal and external features play different roles in face identification (Ellis et al., 1979; De Haan & Hay, 1986; Young et al., 1985; Hines et al., 1987; Bruce et al., 1999; Hancock et al., 2000; Jarudi & Sinha, 2003; Frowd et al., 2007). However, in all these studies internal and external features were presented in isolation. If a task guarantees that the observer selectively focuses either the internal or the external features of complete face stimuli, the contributions of both types of features can be examined, and their temporal unfolding can be assessed in the natural perceptual context.

Unfortunately, there are no measurements at the time allowing to judge the exact times scales of processing internal and external features in the stream of face perception. If it turned out in such measurements that external features are processed fast and show just a marginal FIE, while internal features are processed slower but show a pronounced FIE, then this would be evidence that processing of sets of features conveying a substantial proportion of configural information is indeed slower than processing of sets of features with relatively poor configural content. Since the set of external features includes face outline, or global form, it would not be too surprising if the time scale of this class of features closely resembles the time scale found for object categorisation (Grill-Spector & Kanwisher, 2005). In order to contribute new findings about the time courses of configural and featural modes of processing in face perception we designed the present study, which aimed at measuring the timing of internal and external features in their natural facial context.

5.3 Methods

5.3.1 Experimental outline

Two experiments with identical stimulus material, but different instruction (task) were executed, drawing the observer's attention either to internal or to external features. Experimental control precluded that both types of experiments could be done by focusing only one class of features, ignoring the other. In both experiments a Same/Different task was used, instructing the subjects to compare two subsequent stimulus frames and to decide whether both were same or different with respect to the class of facial features to be attended. In each of both experiments we varied degree of congruency, stimulus orientation, and exposure duration.

Task. In the first experiment (task A: "Same-Internal") subjects were instructed to judge two facial stimuli as *same* when their internal features were congruent, and *different* otherwise. In the second experiment (task B: "Same-External") same-trials were defined as facial congruency in external features. Different subjects participated in both tasks.

Degree of congruency. In each task, facial feature congruency was realized in two degrees, *total* congruency/incongruency (IDE) and *featural* congruency/ incongruency (SISE). In *total* congruency/incongruency, two subsequent face stimuli were either identical faces (T1) or totally different faces (T4). Trials of type T1 correspond to a *same* response, trials of type T4 to a *different* response, independent of task (see Figure 5.1 for the relation of degree of congruency and trial type). In *featural* congruency/incongruency, a trial sequence could contain two faces which were same in internal but different in external features (T2) or same in external but different in internal features (T3). Trial types T2 and T3 correspond to (*same/different*) response alternatives depending on task: In task A

("Same-Internal"), T2, T3 correspond to *same/different* responses, while in task B ("Same-External") this assignment is reversed: T2, T3 correspond to *different/same* responses. Hence, featural congruency/incongruency is difficult than total congruency/incongruency, since there the correct response alternatives depend on the class of features to be attended. Degree of congruency can be looked at with respect to the consistency of the two types of features for the judgement. In IDE, the class of features which is not to be attended provides a consistent facial context for the target feature class, both being same for the *same* response category, and both being different when *different* is the correct response.

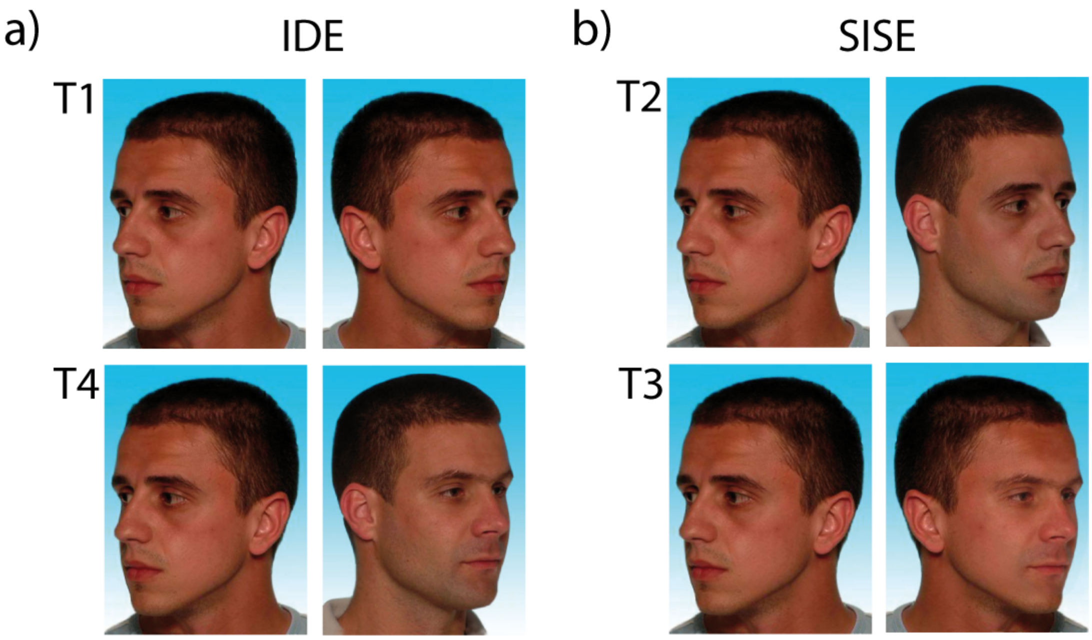

Figure 5.1: Illustration of the two degrees of the facial feature congruence, IDE (a) and SISE (b). In trials of type T1 the same faces were shown, in trials of type T4 the two faces were completely different. T1 and T4 trials form the IDE condition. In T2 trials the two faces had the same internal features, but differ in external features. In T3 trials the two faces were different in internal features, but same in external features. T2 and T3 trials form the SISE condition. Faces were shown in ¾ view, and the face images of a trial sequence were leftward and rightward mirrored examples in order to preclude pixel matching strategies.

In SISE, the features that are not task relevant provide a conflicting facial context, being different when the target features are same, and same when the target features are different.

Orientation. Facial stimuli were presented upright, in their natural orientation, or both stimuli of a sequence were inverted (180^o rotation).

Duration. Six different durations, $D = \{51, 119, 221, 340, 442, 629\}$ ms, were used for stimulus presentation in order to span a wide range of presentation times, ranging from brief timings, precluding saccades and serial scan, up to relaxed timings allowing for detailed and part by part scrutiny of the images. With these factors a 2 (Task) × 2 (Congruency) × 2 (Orientation) × 6 (Duration) factorial design with 48 conditions results, with task as a grouping factor and the remainder as repeated measurement factors[1].

5.3.2 Stimuli

Photographs of four male face models were used as templates for stimulus construction. These were full-color ¾ view photographs of the left face side captured in a photo studio under controlled lighting conditions, and using the same background for all photographs.[2] None of the models was wearing glasses, jewelry or had a beard. Their haircut was comparably similar with short hairstyle (see Figure 5.2).

[1] It is important to note that in each task all four types of trials were presented, and were randomly intermixed. By doing this we precluded subjects from trying to focus just one class of features, ignoring the other, and we guaranteed that the instruction of attending the feature class of interest was obeyed. Looking at the trial examples in Figure 5.2 it comes to mind that a subject might simply pay attention to external features, e.g. to hairs or face outline. In task B he/she says *same* if these features coincide in the two face stimuli of a trail, and *different* otherwise. In task A he/she also focuses external features, but responds with revised assignment to response category, saying *same* if these features differ among the two face stimuli or a trail, and *different* if they coincide. However, since there are also trials T1 and T4, such a strategy would not be effective, and would lead to chance performance. With the acoustical trial by trial feedback, the inefficiency of ignoring the instruction would soon be realised by observer.

The original images were manipulated with Adobe Photoshop in order to construct sample stimuli with defined combinations of internal and external features. Internal features were cut out with comparable tracing lines and placed on the second (template) face, based on the position of the internal features. As a scheme for stimulus construction, we used a **F**acial **F**eature **M**atrix (FFM) with the following structure

$$FMM = \begin{pmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{pmatrix} \quad (1)$$

In (1) an entry F_{ij} denotes a face with internal features of face i and external features of face j. Thus, to form stimulus pairs for "Same-Internal" trials (T2) one has to pair two different faces of a line, which result in $4 \cdot \binom{4}{2}$ possible combinations. Forming "Same-External" stimulus pairs (T3) is done by choosing pairwise combinations from columns, which is also possible in 24 different ways. There are 120 ways of forming "Totally Different" pairs (T4) and 16 ways to build "Identical" pairs (T1). All faces contained in the FFM were mirrored, resulting in a FFM_{left} and a FFM_{right}. Thus leftward and rightward instances of each feature combination were available for selection in a trial sequence.

[2] We employed ¾ view photographs in left and right side perspective, the latter obtained by mirroring the original left face side photographs. Employing both perspectives of the canonical ¾ view has the advantage that pixel region matching strategies for comparing the two faces cannot be employed, and the observer is left to rely on the true comparison of the facial features (see Figure 5.1). The ¾ view has further advantages, e.g. being better identified and generalized to other views than frontal or profile view (for a review see Liu & Chaudhuri, 2002)

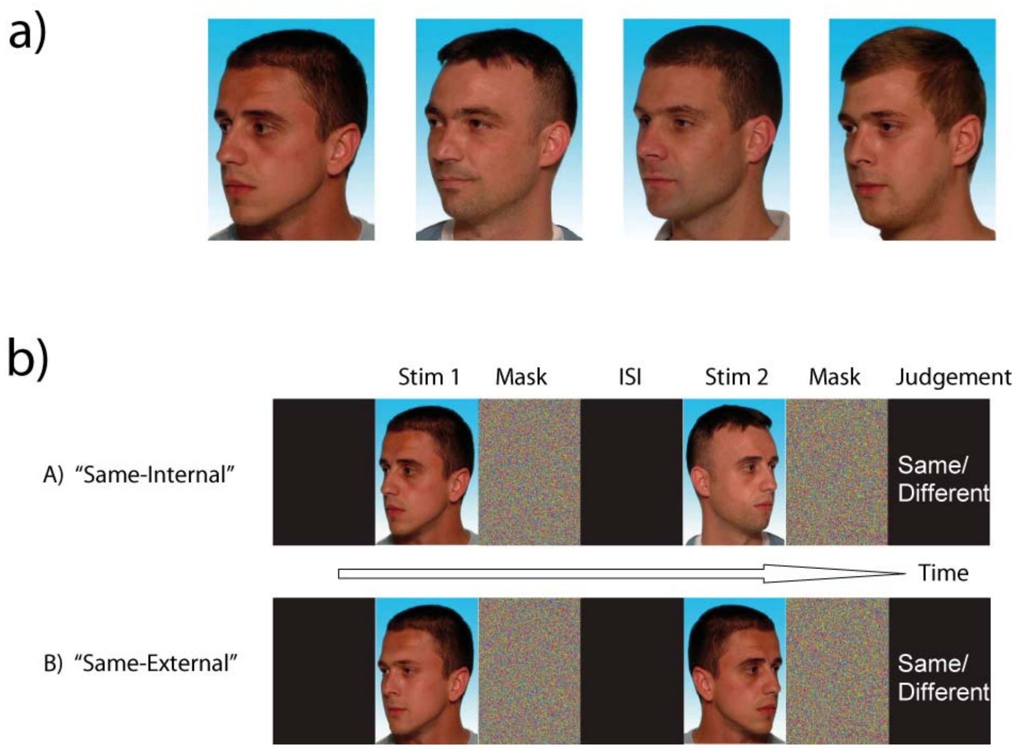

Figure 5.2: The four original faces used for stimulus construction (a) and structure of trial (b). The four original faces form the diagonal entries of FMM (1). The upper panel in b) shows a sample sequence of a "Same-Internal" trial (T2), where the internal features of the two subsequent faces coincide. The lower panel shows an example of a "Same-External" trial (T3).

5.3.3 Performance measures

The experiments aimed at determining the proportion of correct judgements as a function of exposure duration. In order to obtain proportion correct rates free of a possible response bias, they were calculated from correct same *and* correct different judgments. In the *total* congruency/inconcruency condition, the rates were calculated from correct responses to (T1, T4) trials, and in the *featural* congruency/inconcruency condition from correct responses to (T2, T3) trials. Since each trial type was realized in 16 replications, each proportion correct datum rests on $n = 32$ trials.

Since proportions follow a binomial distribution which is approximated by a normal distribution if $np(1-p)>9$, problems with possibly hurt distribution prerequisites may arise in statistical testing with a proportion correct measure, particularly in the vicinity of perfect performance, having p close to 1. To avoid this we transformed the data to d' values, assuming the equal variance case as a convenient quantile transformation.[3]

5.3.4 Subjects

23 subjects participated in task A and 25 in task B. They were undergraduate students, about 20% were male and 80% female. All subjects had normal or corrected to normal vision. The students had no former psychophysical experience, were paid and not informed about the purpose of the experiment.

5.3.5 Apparatus

The experiment was executed with Inquisit 2.0 runtime units. Patterns were displayed on NEC Spectra View 2090 TFT displays in 1280 × 1024 resolution at a refresh rate of 60 Hz. Screen mean luminance L_0 was 75 cd/m^2 at a michelson contrast of $(L_{max} - L_{min})/(L_{max} + L_{min}) = 0.98$, so the background was practically dark (about 1.4 cd/m^2, measured with a Cambridge Research Systems ColorCAL colorimeter). No gamma correction was used. The room was darkened so that the ambient illumination approximately matched the illumination on the screen to a fair degree. Patterns were viewed binocularly at a distance of 70 cm. Stimulus patterns and masks subtended 300 × 400 pixels (width × height), which corresponds to 12 × 15 cm of the screen, or 9.65° × 12° measured in degree of

[3] The two assumptions underlying the d' transformation, equal variance and normal distribution, cannot be directly proven, but the observation of normally distributed residuals in a factorial design may be regarded as consistency proof, since normally distributed residuals are implied by the d' transformations (see Results section). We like to stress that the d' transformation was merely used as a convenient quantile transformation.

visual angle at 70 cm viewing distance. Subjects used a distance marker but no chin rest. They gave responses on an external numeric key-pad, and wore light headphones for acoustical feedback.

5.4 Procedure

5.4.1 Psychophysical task and structure of a trial

A Same/Different forced choice task was used. In task A subjects were instructed to indicate whether the two faces shown in a trial were same with respect to internal features, and in task B whether they agreed in external features. The temporal order of events in a trial sequence was: fixation mark (300 ms) - blank (100 ms) - 1st stimulus frame (Duration) - mask (350 ms) - blank (200 ms) - 2nd stimulus frame (Duration) - mask (350 ms) - blank frame until response. Masking of the stimulus frames was done with spatial noise patterns with a grain resolution of 3 pixels. Acoustical trial by trial feedback was provided about correctness a brief tone signal.

Each subject participated in only one of both tasks. According to the design (see above), each subject had to go through 24 conditions. The measurement for each condition comprised 16 same and 16 different trials, resulting in $24 \times 32 = 768$ trials. These were shuffled and assigned to a randomly ordered measurement list. This list was then subdivided into three blocks with 256 trials, each lasting about 12 minutes. The three blocks were administered to each subject in 3 sessions within two consecutive days.

Proper duration parameters for the 6 durations were found in pilot measurements with some test subjects prior to the main experiments. Further, each subject was made familiar with the task by going through 8 minutes of randomly selected probe trials in order to ensure that the instruction was understood and could be put into practice.

5. 5 Results

5.5.1 Proportion correct as a function of exposure duration

Figure 5.3 shows proportion correct as a function of exposure duration for all experimental conditions. Data are between subjects means, shown with their 95% confidence limits. The proportion correct data were fitted with psychometric curves of exponential form

$$P(t) = 0.5 + b\left(1 - \exp\left(-a\left(t - t_0\right)\right)\right) \qquad (2)$$

having b is an amplitude parameter, a as the shape parameter controling steepness, and t_0 as the location parameter. Parameters were estimated with a least squares criterion using the Levenberg-Marquardt algorithm (Press, Teukolsky, Flamery, & Vetterling, 1996).

Task	*Congruency*	*Orientation*	b	a	t_0	P_{ha}	$t_{.75}$	t_{ha}	t_e
Internal	IDE	Upright	0.449	0.021	17.1	0.724	55.8	50.0	189.6
Internal	IDE	Inverted	0.390	0.017	40.3	0.695	99.1	80.1	229.9
Internal	SISE	Upright	0.340	0.012	52.0	0.670	167.4	112.1	290.5
Internal	SISE	Inverted	0.309	0.007	54.8	0.655	282.1	150.0	356.7
External	IDE	Upright	0.423	0.053	40.2	0.711	57.1	53.3	125.3
Internal	IDE	Inverted	0.409	0.025	32.7	0.705	71.0	60.8	182.8
Internal	SISE	Upright	0.370	0.047	33.2	0.685	57.1	47.9	123.3
Internal	SISE	Inverted	0.381	0.021	33.9	0.690	85.1	67.1	199.8

Table 5.1: Parameters of the psychometric curves for the exponential model (2), and extrapolated duration thresholds. The table shows amplitude parameter, b, scale parameter a, shift parameter t_0, the proportion correct rate for a half-amplitude criterion P_{ha}, and the duration thresholds for the 0.75 proportion correct rate $t_{.75}$, for the half-amplitude criterion t_{ha}, and for the saturation criterion of $\varepsilon = 2.5 \cdot 10^{-4}$, t_ε.

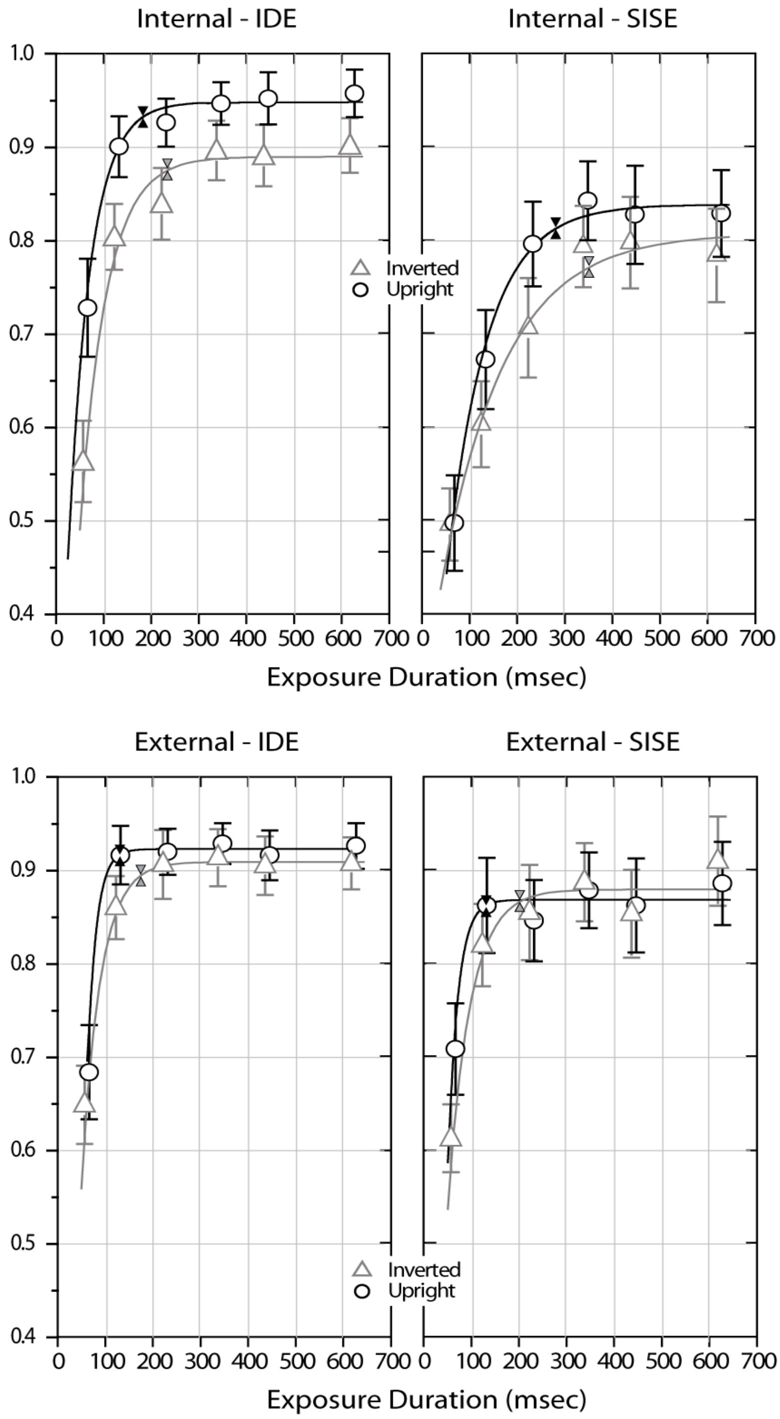

Figure 5.3: Mean proportion correct rates as a function of exposure duration for task A (upper panel) and B (lower panel). The left panel shows the data for IDE condition, and the right panel for the SISE condition. Circles represent performance for upright, grey triangles for inverted faces. The smooth lines are exponential distribution functions (2), with best fitting parameters for the least squares criterion. Small double triangle symbols mart the saturation points of these functions. Error bars denote 95% confidence limits, based on the standard error of the mean for between subject variation calculated for each cell.

For all conditions the model fit was very good, with a ratio of explained to total variation larger than 96%. The parameters obtained from this procedure are listed in Table 5.1.

Inspection of the curves shows that task, degree of congruency and orientation have strong modulating effects, affecting steepness and the saturation level of performance. Generally, psychometric curves enable to determine the exposure durations that are necessary to reach a constant level of performance in each experimental condition. Most convenient are 0.75 proportion correct thresholds. However, in our data the saturation levels differ among conditions. Therefore, absolute performance thresholds should be interpreted with care. To get a complete picture we also considered half amplitude thresholds and saturation points.

5.5.2 Temporal duration thresholds

Since the model (2) describes the proportion correct data fairly good, it is apt for threshold extrapolation. We used (2) to extrapolate 0.75 thresholds, defined as the durations where a proportion correct of 0.75 is reached, and half amplitude thresholds, where performance reaches a level of $0.5 + b/2$. Additionally, saturation points were determined, defined as the durations where the first derivatives of (2) reach a value of $\varepsilon = 2.5 \cdot 10^{-4}$, indicating that the psychometric curves (2) have settled to an almost constant value. These three types of critical durations are shown in Figure 5.4, and their values are listed in Table 5.1. Additionally, saturation points are marked in the psychometric curves (see Figure 5.3) with small double triangle symbols. The differences in the temporal thresholds for the two tasks are striking. For matching external features both the 0.75 and the half amplitude thresholds remain below 85 ms, are nearby, and are not strongly modulated by the degree of congruency. For matching internal features these thresholds strongly increase in featural congruency/incongruency (SISE) compared to total congruency/incongruency (IDE). The saturation points reflect this tendency even more pronounced. For matching external features the time necessary to come close to saturated performance in each experimental condition is only modulated by stimulus orientation, but not by degree of congruency/incongruency. For matching internal features these times are modulated by both factors, but stronger by the degree of congruency/ incongruency.

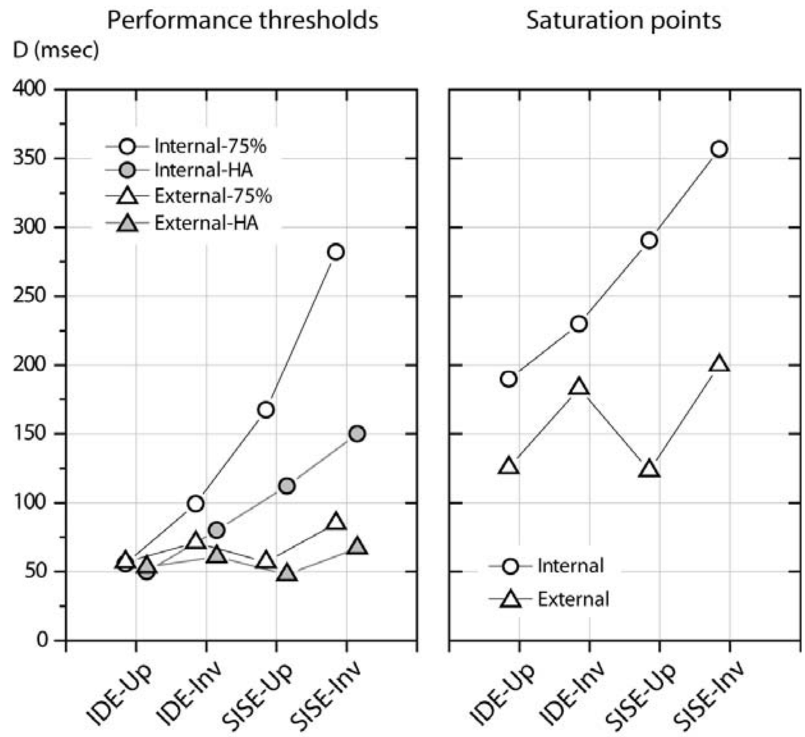

Figure 5.4: Temporal duration thresholds, extrapolated from psychometric curves. The left panel shows 0.75 proportion correct thresholds (open symbols) and half amplitude thresholds (grey symbols), the right panel shows saturation points for a saturation criterion of $\varepsilon = 2.5 \cdot 10^{-4}$.

The saturation points are also illustrated by double triangle markers directly in the psychometric function graphs shown in Figure 5.3. For matching external features (see lower panel of Figure 5.3), psychometric curves are relatively steep, and reflect high performance within the first 120-200 ms. Comparing the locations of the triangle markers shows that these are at practically the same locations for IDE (left panel) and SISE (right panel), but are rightward shifted for inverted faces relative to upright faces to a similar amount in both congruency conditions. For matching internal features (see upper panel of Figure 5.3) psychometric curves are generally flatter. In this task, the saturation points are shifted rightward on the temporal scale for SISE relative to IDE, and this shift is stronger than the rightward shift observed for inverted (grey markers) relative to upright face stimuli (black markers).

Taken together, analysis of psychometric curves shows that matching of external features is relatively fast, reaching high performance levels within the first 120 to 200 ms. Matching of internal features is slower, evolves more gradually in time. Performance is more subject to higher task demands as induced by featural congruency/incongruency, compared to total congruency/ incongruency. Saturated performance is reached at the earliest after about 200 ms in IDE, and after about 300 ms in SISE.

5.5.3 Testing matching performance levels

As the psychometric functions of exposure duration indicate, matching performance is strongly modulated by task, degree of congruency/incongruency and orientation. To avoid problems with hurt assumptions in statistical testing we calculated d' data from the detection rates. Figure 5.5 shows mean d' data as a function of exposure duration. These functions take, in principle, the same course as the proportion correct functions shown in Figure 5.3. As the proportion correct data, the d' data can also be approximated with exponential distribution functions, but do not require the guessing constant in (2). In Figure 5.5 the least squares fits for this model type are shown as solid lines.

For statistical testing the d' data were fed into ANOVA routines. F - tests with Greenhouse-Geisser corrected degrees of freedom were calculated, precluding progressive testing results due to hurt sphericity assumptions. Normality of within-cell residuals was assessed for all 24 cells with a q-q plot method (Johnson & Wichern, 2003, p. 155ff), showing an excellent fit of observed and predicted residuals with correlation coefficients larger than 0.94. No systematic deviations at the tails, indicating skewness, were observed. Table 5.2 shows the results of the F -tests. With the exception of two high order interactions, all effects are significant. Lack of significance of the congruency × orientation × duration interaction means that the curves shown in Figure 5.5 are parallel when collapsed over both tasks.

The high significance of the interaction of highest order (congruency × orientation × duration × task) is due to the much higher steepness of the curves for matching external compared to matching internal features, as already discussed with the psychometric curve data. In order to explore the temporal dependency we illustrate and discuss the effects of the main factors as a function of exposure duration.

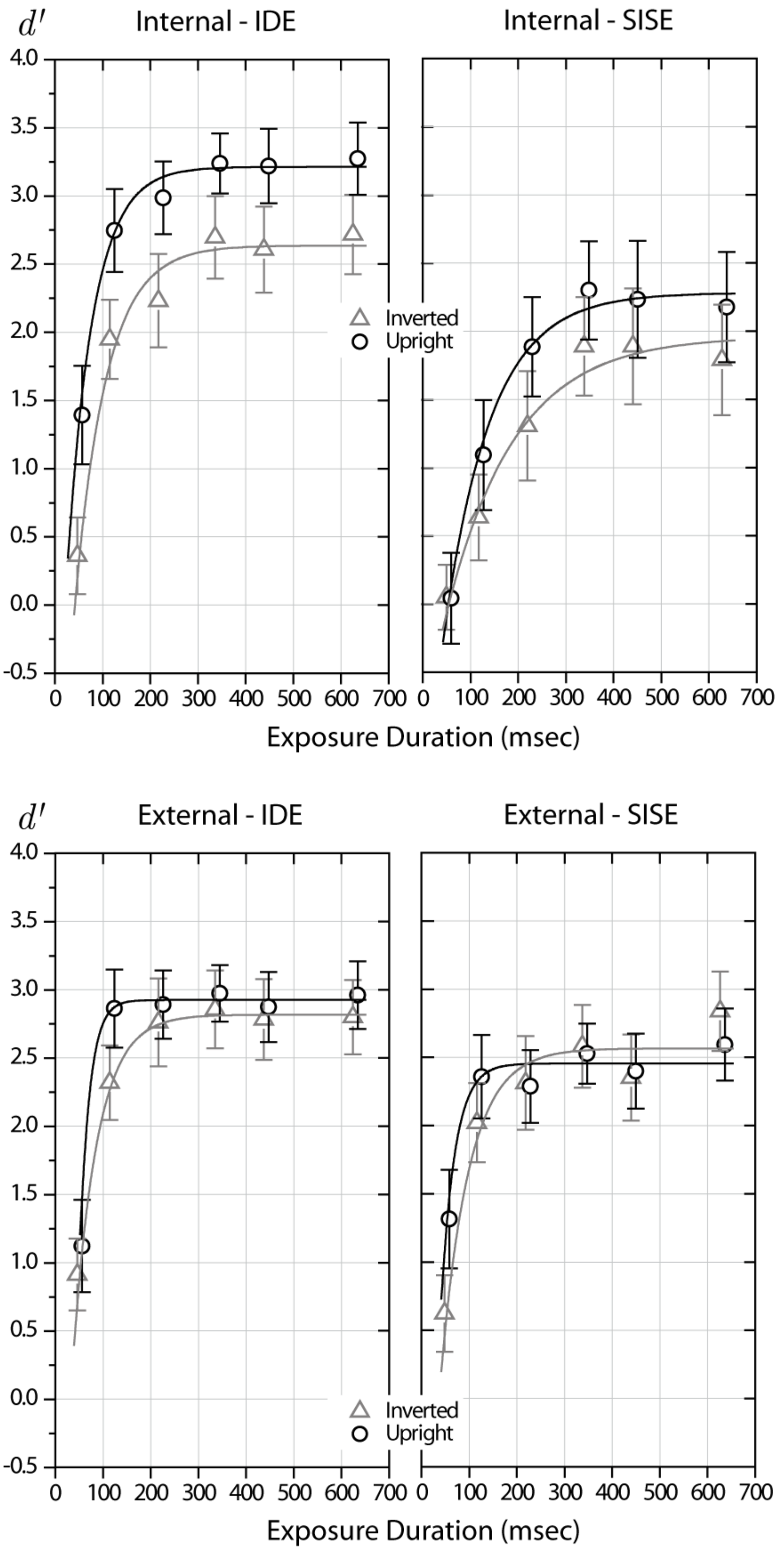

Figure 5.5: Mean d' data as a function of exposure duration. Arrangement of the data panels and conventions are the same as in Figure 5.3.

Source of Variation	*SS*	*df*	$\hat{\sigma}^2$	*F*	*p*	df_1	df_2	$\tilde{P}$	η^2
Task (A)	49.759	1	49.759	6.922	0.012				
Error	330.646	46	7.188						
Congruency (B)	128.808	1	128.808	139.098	0.000	1	46	0.000	1.000
A x B	35.743	1	35.743	38.598	0.000	1	46	0.000	1.000
B x subjects	42.597	46	0.926						
Orientation (C)	34.872	1	34.872	101.346	0.000	1	46	0.000	1.000
A x C	9.44	1	9.44	27.435	0.000	1	46	0.000	1.000
C x subjects	15.828	46	0.344						
Duration (D)	514.389	5	102.878	235.773	0.000	4.09	188.19	0.000	4.090
A x D	13.948	5	2.79	6.393	0.000	4.09	188.19	0.000	4.090
D x subjects	100.359	230	0.436						
B x C	3.712	1	3.712	16.703	0.000	1	46	0.000	1.000
A x B x C	1.497	1	1.497	6.734	0.013	1	46	0.013	1.000
B x C x subjects	10.224	46	0.222						
B x D	7.199	5	1.44	5.13	0.000	3.72	171.25	0.001	3.720
A x B x D	4.421	5	0.884	3.15	0.009	3.72	171.25	0.018	3.720
B x D x subjects	64.554	230	0.281						
C x D	4.538	5	0.908	3.014	0.012	3.66	168.46	0.023	3.660
A x C x D	2.304	5	0.461	1.53	0.181	3.66	168.46	0.200	3.660
C x D x subjects	69.26	230	0.301						
B x C x D	0.257	5	0.051	0.291	0.918	4.51	207.64	0.903	4.510
A x B x C x D	5.971	5	1.194	6.759	0.000	4.51	207.64	0.000	4.510
B x C x D x subjects	40.637	230	0.177			1	46	0.000	1.000

Table 5.2: ANOVA results for the d' data shown in Figure 5.5. The table shows source of variation, sum of squares, *SS* degrees of freedom, *df*, variance estimate $\hat{\sigma}^2$, *F*-ratio, *F*, and significance level, *p*. The next three columns show the results of the Greenhouse-Geisser corrected *F*-tests compensating for lack of sphericity: corrected degrees of freedom for denominator df_1, and nominator df_2, and the resulting significance level, $\tilde{P}$, for the *F*-ratio with the corrected degrees of freedom. The last column holds the ratio of explained of total variation, η^2.

5.5.4 The face inversion effect (FIE)

The effect of face inversion is captured by the difference in the d' measure obtained for upright and inverted facial stimuli. The overall d' difference for both stimulus orientations is strongly significant, but orientation variation explains just 3.1% of the total d' variance, a substantially smaller proportion than explained by task and congruency/incongruency (see Table 5.2). For analysing the temporal dependency of the FIE we calculated the difference $\Delta d' = d'_{Up} - d'_{Inv}$ on the level of individual subject data, for each duration, congruency condition, and for both tasks. If the expected values for upright and inverted stimuli are the same in each experimental condition, the expected value of $\Delta d'$ is zero. Hence, a conservative assessment of the FIE is proving whether zero lies within the confidence limits spanned about the actual mean difference $\Delta d'$. This confidence interval is determined by the standard error of $\Delta d'$ for each j-th experimental condition, i.e.

$$CI_j = \overline{\Delta d'_j} \pm t_{(N-1;0.975)} \sqrt{\frac{\hat{\sigma}_j^2}{N}} \qquad (3)$$

defines this confidence interval in the j-th experimental condition. Here, $\hat{\sigma}_j^2 = \sum_i \left(\Delta d'_{ij} - \overline{\Delta d'_j}\right)^2 / (N-1)$ is the unbiased estimate of the population variance of the difference $\Delta d'$ in condition j, and $t_{(N-1;\,0.975)}$ is the critical t- quantile for a 5% alpha level (two-tailed). For convenience, we pooled the variance estimates for IDE and SISE conditions at each duration and calculated joint confidence intervals for the mean difference, spanned around 0, for each task. These intervals are shown as grey shaded areas in Figure 5.6. Mean differences lying outside this area deviate significantly from zero. The mean difference data for the IDE condition in task A are above this area at all timings, indicating a strong face inversion effect for matching internal features with totally congruent/incongruent facial stimuli. In the more difficult SISE condition there is a much smaller FIE, which fails significance at two durations. For matching external features a FIE is absent for exposure durations larger than about 200 ms, but for briefer timings smaller than 120 ms the effect also exists, with a strength of about half a d' unit, indicating that the effect vanishes in the time window between 120 ms and 200 ms.

To further explore the dependency of the FIE on task and degree of congruency/incongruency we analysed the difference data with ANOVA routines, having task as grouping factor and degree of congruency/incongruency and duration as repeated measurement factors.

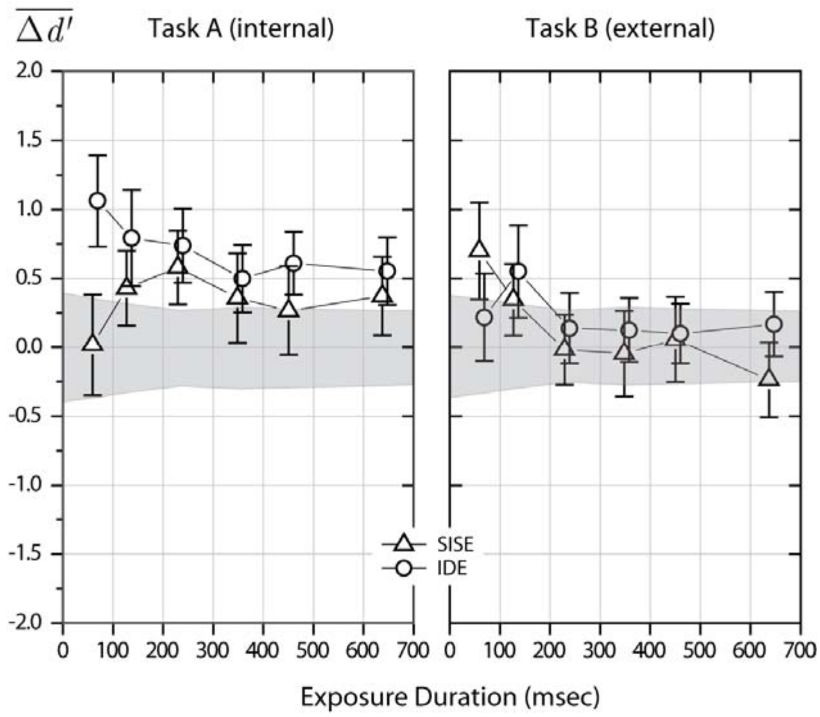

Figure 5.6: Difference measure $\Delta d'$ for the performance with upright and inverted stimuli as an indicator of the strength of the face inversion effect (FIE). Mean difference data are shown for matching internal features (task A, left) and matching external features (task B, right). Error bars denote 95% confidence limits, based on the standard error of the mean for between subject variation in each condition. The grey shaded area marks the confidence limits spanned around 0, based on a standard error pooled across SISE and IDE data at each exposure duration.

The results prove that the FIE is stronger for matching internal than for matching external features ($\eta^2 = 0.374$, $F(1,46) = 27.436$, $p < .001$), and that it is stronger for total (IDE) compared to featural congruency/incongruency (SISE) ($\eta^2 = 0.027$, $F(1,46) = 16.703$, $p < .001$). Also the general decline of the FIE with increasing exposure duration is significant ($\eta^2 = 0.033$, $F(5,230) = 3.014$, $p < .01$).

5.5.5 The effects of total (IDE) and partial (SISE) stimulus congruency/incongruency

The effects of the two degrees of congruency/incongruency can be analysed by looking at the difference in the d' measure obtained for both conditions. Also this overall difference is highly significant, and explains 11.6% of the total d' variation (see Table 5.2). As done for exploring the effect od stimulus inversion, we further analysed the difference measure $\Delta d' = d'_{IDE} - d'_{SISE}$ for each duration and stimulus orientation in both tasks. Again, confidence limits for the mean differences $\Delta d'$ were calculated, following the same procedures as used for analysing the FIE (see above).

The grey shaded areas in Figure 5.7 mark the confidence intervals for the mean difference spanned around zero on the basis of a joint variance estimate for upright and inverted stimulus orientation at each duration. The data of Figure 5.7 indicate that there is a performance advantage for IDE compared to SISE in both tasks, but this advantage is much stronger for matching internal than for matching external features. Beyond 200 ms the advantage for the IDE condition is constant, no more depending on exposure duration. To have this analysed further, we fed the difference measure in ANOVA routines with task as grouping factor and orientation and duration as repeated measurement factors. This analysis revealed that the IDE/SISE effect is much stronger in task A (matching internal features)

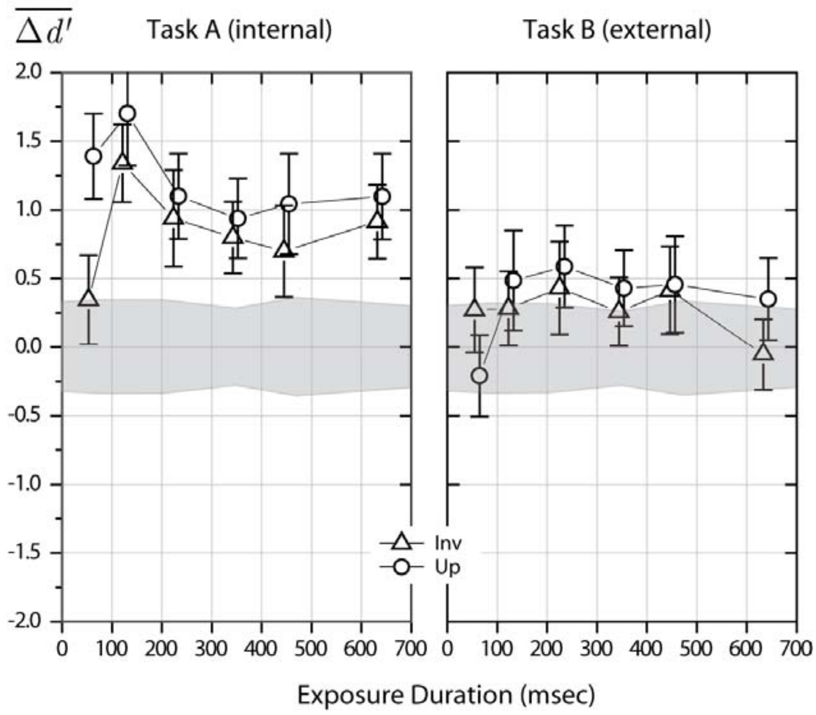

Figure 5.7: Difference measure $\Delta d'$ for the matching performance difference achieved in IDE and SISE congruency conditions. Mean difference data are shown for matching internal features (task A, left) and matching external features (task B, right). Error bars denote 95% confidence limits, based on the standard error of the mean for between subject variation in each condition. The grey shaded area marks the confidence limits spanned around 0, calculated with a standard error pooled across the data for upright and inverted stimulus orientations at each exposure duration.

than in task B (matching external features, (η^2 = 0.456, $F(1,46)$ = 38.598, p < .001), and for upright compared to inverted stimuli (η^2 = 0.027, $F(1,46)$ = 16.703, $p < .001$). There is also an effect of duration (η^2 = 0.052, F(5,230) = 5.130, $p <$ .001), but this effect just reflects a high variability at the first two durations.

5.5.6 The effects of instruction

The effects of the two different tasks cannot be analyzed in the same way as done for stimulus orientation and congruency/inconcruency condition, since task is a grouping factor, which means that the two tasks were executed by different subjects. Hence a difference measure cannot be defined on the d' data for each subject in each condition. Nevertheless, a d' difference can be considered on the level of the means, defining $\Delta\overline{d'}_{BA} = \overline{d'_B} - \overline{d'_A}$. This difference follows a t-distribution with df = $N_B + N_A$ - 2, and has standard error $\hat{\sigma}_e = \sqrt{\hat{\sigma}_B^2 / N_B + \hat{\sigma}_A^2 / N_A}$, allowing to calculate a confidence interval $CL_{BA} = \Delta\overline{d'}_{BA} \pm t_{(N_B+N_A-2;1-\alpha/2)}\hat{\sigma}_e$ for each congruency/incongruency condition and each stimulus orientation. The differences of the task means and their confidence limits are shown in Figure 5.8. Again, a joint confidence interval for upright and inverted stimuli was calculated (see grey shaded area in Figure 5.8).

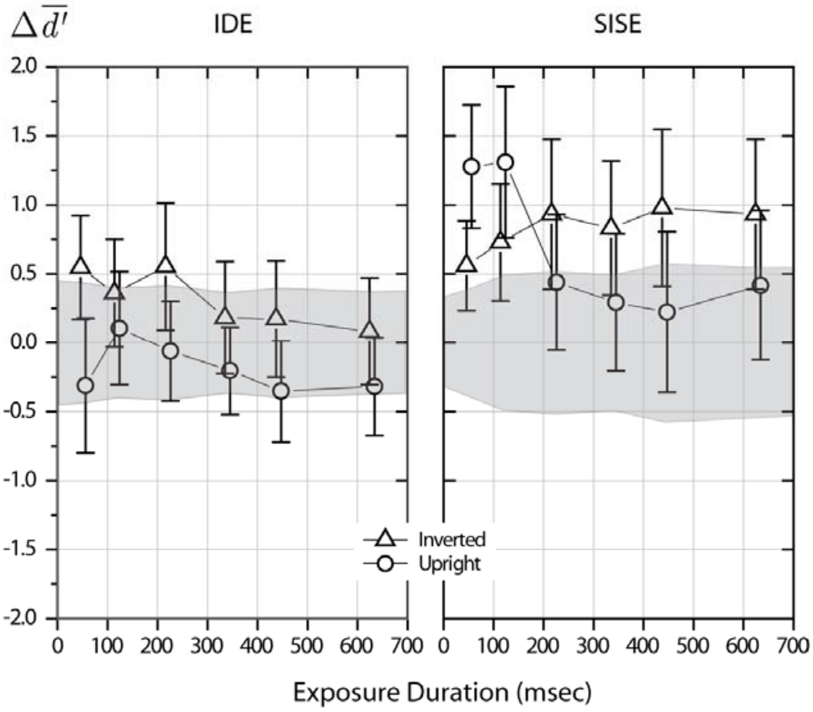

Figure 5.8: Differences of means in the d' measure for the two task instructions B and A, matching of external and matching of internal features. Data points are differences of the task means. Error bars denote 95% confidence limits for the differences of means, based on the standard errors of the means in the source samples of each condition. The grey shaded area marks the confidence limits spanned around 0, calculated with a standard error pooled across the data for upright and inverted stimulus orientations at each exposure duration.

The overall task effect is strongly significant, and is, after duration, the second largest source of variation, explaining 13.1% of the total d' variance (see Table 5.2). The task effect, stronger in SISE than in IDE ($t = 4.574$, $df = 22$, $p < .001$), is significant at all durations only in SISE with inverted stimuli. A look at Figure 5.5 confirms that matching internal features in SISE with inverted stimuli leads to the worst performance of all conditions, while matching external features does not suffer much from stimulus inversion in SISE. A further striking property of the data is that there is no performance advantage for matching external features with upright stimuli in IDE, indicating that the observers have access to an additional source of information. However, this trend is not significant. Moreover, there is a tendency of better performance in matching of external features at brief timings, since the difference of means tend to fall with exposure duration.

A striking property of all the three difference measures is that they settle to an almost constant level for exposure durations of beyond 200 ms, and are different within the first 120 ms of processing. This indicates that the interval of 120 - 200 ms is critical for a new source of information which enters the matching process, exerting their full influence after 200 ms have passed.

5.6 Discussion

In a face matching task we measured proportion correct as functions of exposure duration and determined how these functions are modulated by image inversion, the congruency of facial feature context, and by the types of features to be attended. Results reveal new aspects for the distinction of internal and external features, and indicate that both types of features involve different kinds of information in the stream of processing.

Timing differences among internal and external features. Analysis of duration thresholds showed that matching of external features can be done at an accuracy level of 75% already within the first 85 ms, and is robust against higher task demands. In contrast, matching of internal features at the same levels of accuracy takes more time, rising up to durations of about 280 ms when internal features are to be compared in inverted presentation with conflicting facial context (SISE). Albeit the difference in the saturation level of performance, the temporal course and slopes of the psychometric curves for IDE and SISE are very close for external features, while for internal features the SISE curves are much flatter than the IDE curves. This indicates that extraction of external features occurs automatic, and does not require allocation of additional resources with increasing task demands. For internal features, however, feature extraction in a conflicting facial feature context requires resources, and, therefore, time.

Early and later steps of processing facial features. The psychometric curves for upright stimuli in IDE are very close for internal and external features, with higher saturation levels reached for internal features after about 200 ms. Within the first 120 ms the curves are practically not distinguished (see Figure 5.3). The corresponding curve for internal features in SISE shows that internal feature matching in conflicting facial contexts is not really possible within the first 120 ms, being at chance level for 51 ms and at 67% at 119 msec. On the other hand, the psychometric curve for external features in SISE is already at its saturation level at 119 ms. This shows a striking performance asymmetry for internal and external features within the first 120 ms of processing. While external features can be selectively attended in face matching, independent of the state of congruency of the internal features in the two faces to be compared, extraction and selective comparison of internal features is seriously hampered by an incongruent external feature context. Since, apparently, matching based on just internal features is poor within the first 120 ms, the high matching performance in the IDE-internal condition cannot be based on the representation of internal feature sets.

Instead, poor matching of internal features in conflicting facial contexts (SISE) and same performance with internal and external features in congruent facial contexts (IDE) indicates that performance with internal and external features has the same origin in the first 120 ms of a IDE condition. Since extraction of isolated internal features and establishment of their spatial relationships to other facial parts is not yet accomplished in this brief time interval, the observer can only rely on external features, which provide global shape information, but no fine details and their inner structure. Later, internal features, with spatial relations among each other and to external features, add.

The presumption that internal features and configural information enter with some temporal delay is supported by the asymmetry of the feature context congruence effect (IDE/SISE effect) for internal and external features, and its time course (see Figure 5.7)[4]. In the first 120 ms of processing matching performance based on purely internal feature sets is poor (SISE-internal), but quite good when the observer can rely on external features instead (IDE-internal), leading to a pronounced difference of both conditions within this time interval (see left panel of Figure 5.7). After this time, internal features and configural information become available, enabling the observer to exploit spatial relations within internal features, but, more important, also to external features. Comparing the IDE/SISE effect for both tasks shows that, after the first 120 ms, agreement/disagreement in both types of features leads to performance improvement of more than one d' unit when the observer focuses internal features, while this effect is less than half this size when he/she focuses external features (see Figure 5.7 and related statistical testing). The plus of 0.5 d' units in the IDE/SISE effect indicates that perception of internal features is strongly modulated by an additional source of information that becomes available with the coincidence of both types of facial features.

[4] This is also supported by the time course of the face inversion effect (see below).

This source of information is the spatial relationships of internal and external features, enabling the comparison of complete 'wholes'. For an observer who focuses global shape, or face outline, spatial relations are of minor importance. He/she will benefit from the additional coincidence of the inner facial areas, but not from the impression of increased coherency, which is more important to an observer who inspects the inner region of a face.

The claim that the representation of internal features is just rudimentary within the first 120 ms of processing, while face outline, hue and global shape information is fully developed within this interval is supported by recent findings, and fits to the scheme of results accumulated in the global/local debate in object perception.[5] Surgase and colleagues (Surgase, Yamane, Ueno, & Kawano, 1999) recorded from face sensitive cells of the macaque temporal cortex, and found that global information, categorizing stimuli as monkey faces, human faces or shapes, was contained in the earliest part of the responses. Fine information about identity or expression was transmitted later, with a latency of 51 ms after the onset of global information transfer. Grill-Spector and Kanwisher (2005) led subjects detect, categorize, and identify objects from briefly presented images, and determined performance thresholds from psychometric curves. They found that when subjects were able to correctly indicate that there was an object contained in an image, they could also indicate its basic level (e.g. a dog, car or flower). Correctly indicating the subordinate level (e.g. porsche, shepherd dog, rose) at the same level of correctness took about 65 ms longer. Most important, the corrected 50% performance thresholds for detection and categorization were found to be at about 50 ms, which indicates that object categorization is rapid, while focal feature extraction allowing to distinguish among class members lasts about double the time.

[5] It should be pointed out that the employed stimulus set may have emphasized contrasting results for internal and external features, since the four basis faces used for stimulus construction are rather similar, lacking single distinctive features (see Figure 5.2). Matching them on the basis of internal features benefits from exploring configural information. If this kind of information comes later in the stream of processing, a pronounced performance difference for a focus on internal and external features is expected.

These results, substantiated by earlier findings of rapid object categorization in visual scenes revealed by EEG (Thorpe, Fize, & Marlot, 1996), indicate that object analysis proceeds from coarse to fine, with outlines and shapes being available first, while details require scrutiny, resulting in longer times (Ahissar & Hochstein, 1993, 1996, 1997). With more abstract stimuli this principle was already formulated three decades ago by Navon (1977), based on a global/local response time asymmetry found for judging global letters made up by local letters, where the subjects' focus was either on the global letters or its local constituents. Also neuropsychological findings suggest that global and local information are processed in parallel by independent modules located in different hemispheres, with the local information having some latency (see Robertson & Lamb, 1991, for a review).

The face inversion effect. In our study we found a pronounced face inversion effect (FIE), which is commonly taken to indicate a loss of configural processing in inverted faces, while upright faces are seen as wholes with an innate structure (see Section 5.2). Closer analysis of this effect substantiates that configural information is not available within the first 120 ms, but enters later, being fully developed after about 200 ms have passed.

As a striking result, our data show that there is a FIE at brief timings, which is independent of the nature of features that are monitored (see Figure 5.6). After the first 120 ms a FIE for external features fades, but resides at a strong level for matching of internal features, and is more pronounced in congruent facial contexts (IDE) than in incongruent ones (SISE). This scheme of results indicates that the FIE within and beyond the first 120 ms of processing has different origins. Within the first 120 ms the effect is "featural", reflecting that comparison of external facial features is hampered for brief timings, since also face outlines have a "usual" orientation as facial stimuli. With more time the featural comparison can be accomplished despite the unusual orientation, and the advantage of upright stimuli vanishes when external features are predominantly monitored (see right panel of Figure 5.6). However, an observer who monitors the inner region of a face benefits from the availability of configural information after 120 ms, which is intact in upright but disrupted in inverted faces. As a result a substantial FIE resides at a constant level independent of further increasing viewing time. The FIE is even more pronounced for congruent external feature contexts (IDE), enabling the observer to additionally exploit the spatial relations to external features, and, henceforth, letting him/her compare integrated wholes (see left panel of of Figure 5.6). Thus the strongly transitive order in the plateau phase of the FIE at all durations of 221 ms and beyond

FIE (IDE, internal) > FIE (SISE, internal) > FIE (IDE & SISE, external)

indicates that a different amount of configural information is effective in the corresponding conditions: spatial relations of all face parts in IDE-internal, spatial relations among internal features in SISE-internal, and no or just a marginal portion of spatial relations when external features are focused.

The majority of studies dedicated to the FIE with internal and external features supports the basic observation that the FIE is larger for internal features, and marginal or absent for external features (Moscovitch & Moscovitch, 2000; Nachson & Shechory, 2002; Veres-Injac & Schwaninger, submitted). However, except in the present study, the timing of the FIE has so far not been addressed.

Different processing systems for internal and external features? With the pronounced timing differences for internal and external facial features, as well as their implications for the different kinds of information available at different moments in time, it is worth to ask whether both types of features are processed by different subsystems.

In an interesting study by Moscovitch & Moscovitch (2000) the authors compared face recognition performance of normal subjects with the performance of a subject with object agnosia, but normal recognition of upright faces (patient CK). While normal subjects could solve a face recognition task with only external features (with the inner face part clipped off), CK's performance was impaired. With inverted whole faces his performance was seriously impaired, while normal subjects still had good recognition rates. These observations indicate that external features play a significant role in face recognition, and that external features also stimulate the object recognition system, which interacts with the face recognition system in a face recognition task. Neuropsychological evidence from patients with prosopagnosia, the counterpart of object agnosia as the selective impairment of face recognition, supports this conclusion. There, recognition of inverted faces is at the level of normal subjects, suggesting that perception of inverted faces is predominantly mediated by the object recognition system (Farah et al., 1995). In fact, some prosopagnosic patients were found to perform worse with upright than with inverted face stimuli (ibid.). Studies using fMRI revealed that the effect of face inversion was an increased response in ventral extrastriate regions that respond preferentially to other classes of specific objects (e.g. houses). In contrast, inversion of houses did not produce a similar effect in face-selective regions (parahippocampal, fusiform and inferior temporal gyri), but led to a reduced activity in the same regions (Haxby et al., 1999).

The timing approach is a possible means to disentangle the contributions of both systems. If the object recognition system operates fast while the face recognition system is enabled with some delay, and if internal features stimulate just the face recognition system while external features stimulate both systems, then we would expect exactly the scheme of results we actually observed. We found that in the first 120 ms IDE-internal, IDE-external and SISE-external conditions showed exactly the same time courses within each orientation, upright and inverted. This indicates that the same system is activated in all these conditions, which preferentially processes outlines and external features. After the first 120 ms we observe different saturation levels, depending on instruction, and a vanishing FIE in task B (matching of external features), but a prevailing FIE in task A (matching of internal features). This indicates that another system comes in, which responds to the configural information of faces.

Corresponding to our psychophysical findings different time scales were found in fMRI and ERP studies of object recognition and face recognition. In object categorization fMRI studies have shown that activation in object tuned areas, and mostly in the LOC, is correlated with masked image presentations at brief timings of 40 ms up to 100 ms (Grill-Spector et al., 2000). ERP studies identified the N170 component, which starts at 130 ms and peaks at 160 ms, being elicited by faces, but not by houses, cars, hands, or furniture. This component has been shown to encode individual face differences, to be modulated by internal features, but not by external features, and also to be strongly sensitive to face inversion (Bentin, Allison, Puce, Perez, & McCarthy, 1996; George, Evans, Fiori, Davidoff, & Renault, 1996; Eimer, 2000; Jacques & Rossion, 2006; Jacques et al., 2007).

Revisiting the claim of Carbon & Leder (2005) that featural information precedes configural information we arrived at the conclusion that face outlines and shapes precede spatial relations, and found evidence for a critical time window between 120 ms - 200 ms in which configural information establishes. Further, we found an early feature selective FIE which vanishes in the course of processing, supporting the author's interpretation of the Thatcher advantage effect at brief timings. However, several open questions remain. First, our interpretation that within the first 120 ms the object recognition system is predominantly involved should be substantiated by further consistency proofs, showing similar processing characteristics in the first 120 ms for faces and objects, and different characteristics afterwards. Particularly, an inversion effect for objects should exist for short exposure durations, and vanish in the further course of time. Second, the critical time window of 120 ms to 200 ms should be further explored, and experiments on the detection of configural changes in faces should be executed with the timing approach. If configural information is not available before 120 ms, but fully established after 200 ms, then these critical timings should also be observed in experiments on matching of spatial relations among face parts. This is forthcoming work.

6. General Discussion

The basic mechanisms involved in early face perception were studied by exploring featural and configural information processing in the context of natural face stimuli. A major aim of all three studies was to contribute do identifying processing paths in face perception activated by internal (eyes, eyebrows, nose, mouth) and external (hair, head and face outline, ears) facial features, and to draw conclusions about configural and featural modes of processing. According to previous studies internal and external features play different roles in face perception (Ellis et al., 1979; De Haan & Hay, 1986; Young et al., 1985; Hines et al., 1987; Bruce et al., 1999; Hancock et al., 2000; Jarudi & Sinha, 2003; Frowd et al., 2007), and, presumably, both types of features activate different visual subsystems (Moscovitch & Moscovitch, 2000). However, external and internal features have so far mostly been studied in isolation, which implies that the conclusions gained from most previous studies do not necessary generalize to the modes of processing used by both types of features in an intact, natural facial context.

The specific contributions of each of both feature types to face perception can be properly examined only within the context of whole faces, where not only featural, but also configural and/or holistical information is available. The whole natural facial context assures that all three kinds of information can, in principle, be activated by each of the two types of facial features, or their interplay, and allows to study the relationship of feature type and the kind of information it predominantly transmits. Further, it allows to study the temporal evolvement of featural, configural and holistic information with time.

In order to reveal the efficacy of the three kinds of information at different moments in time we created facial stimuli by combining internal and external features of four different faces, and devised a task that guaranteed that the observer selectively focused one of both types of features. With this basic experimental paradigm we studied the effects of inversion, viewpoint, congruency of facial context, and exposure duration for both instructions. Focusing on the effects of inversion and viewpoint in the first two studies (Chapter 3 and 4) it was possible to define two distinct processing paths for internal and external facial features, and to draw conclusions about their interaction in producing a holistical facial percept. In Study 3 (Chapter 5) the question whether face processing is predominantly featural or configural was pursued by exploring the timing of external and internal features. Evidence for hypothetical processing stages and their temporal order has been found, based on the observation of a rapid availability of information provided by external features, while information provided by internal features was shown to become perceptually salient at later moments in time.

Particularly, our findings on the effects of inversion and viewpoint for different exposure durations enlarge current knowledge about hypothetical processing paths, and the integration of featural and configural information in face perception. They lead to a clear picture of the different contributions of internal and external features in face perception, the different subsystems that are predominantly activated by both feature types, and the way they interact in the stream of processing in order to produce a holistic face percept. In the following sections the results and conclusions gained from our three studies will be discussed in a global framework, and related to the leading approaches of the field.

6.1 Featural and configural information in face perception

As introduced in Chapter 1, ample evidence exists for at least two main types of information that are transmitted by facial stimuli: featural, defined by the specificities of isolated features, and configural, comprising relations among features and their unique spatial organization.

In Study 1 (Chapter 3) we examined processing of facial features and configurations following two lines of evidence. Firstly, we focused on processing of identical (ID) and different (DF) faces, where the information yielded by internal and external features was congruent (i.e. faces were either completely same or completely different), compared to processing of faces sharing only same internal (SI) or only same external (SE) features, where information yielded by the types of features which were not to be attended was incongruent with the correct response. The results reflect a clear advantage for whole face matching in congruent conditions (ID and DF), compared to incongruent conditions (SI and SE). The congruency effect is in line with the claim of Nachson et al. (1995) that faces are “configurational” in the sense that the various features interact to produce an integrated facial stimulus (Nachson et al. 1995). The whole-to-part advantage in face perception was also obtained in whole-part experiments (Tanaka & Farah, 1993; Tanaka & Sengco, 1997) and experiments with chimetric faces (Hole, 1994; Hole et al., 1999; Young et al., 1987). Corroborating our findings, these studies provide ample evidence for the claim that strong spatial interactions influence perception of features, or halves of faces, in upright orientation.

The second line of evidence regarding whole-to-part face processing in Study 1 followed matching criteria in three different tasks: matching of whole faces (Task 1), matching of external features (Task 2) and matching of internal features (Task 3). When presented in isolation both external and internal features can be quite effective for matching of unfamiliar faces (Ellis at al., 1979; Hines et al., 1987). However, when presented in full facial context, it can be expected that the respective efficacy of each feature class may be differently affected when subjects are required to focus on either one of them while performing the face matching task. In Study 1 no differences in reaction times were obtained in the three tasks employed, suggesting that processing time for whole faces as well as facial features may occur at the same time scale. However, there were significant differences in the accuracy data, where internal features were found to be less accurately matched than whole faces and external features. This indicates that the contribution of external and internal facial features to face matching is not symmetrical, where external features seem to be much easier accessed, and are seemingly more efficient in affecting decisions about face matching (Nachson et al, De Haan & Hay; 1986; Young, et al., 1985; Bruce et al., 1999; Hancock et al., 2000; Jarudi & Sinha, 2003; Frowd et al., 2007).

One possible explanation for the observed asymmetry may derive from the so-called *global precedence phenomenon* (for a review see Robertson & Lamb, 1991), which may also account in face perception. The theory of a global precedence was originally proposed by Navon (1977, 1981), who saw it based on two effects. Using patterns that comprised local letters nested within global letters, he found that reaction times were faster for global than for local letters ("global advantage"). When global and local letters did not coincide (inconsistent condition), reaction times for responding to the local letters were impaired (global interference), but not vice versa. If faces are conceived as hierarchical patterns, then external and internal features correspond, by analogy, to global and local levels, respectively. Indeed, there is evidence that hairline and chin may be processed more globally, whereas features as mouth and nose are processed locally and in feature-by-feature manner (Matthews, 1978; Walker-Smith, Gale & Findlay, 1977). Hence, the mechanisms proposed by the global precedence hypothesis could explain better matching performance for external compared to internal features. However, global precedence cannot explain why the performance advantage is only obvious in accuracy, but not in reaction times.

In Study 2 (Chapter 4), there was a clear advantage in both reaction times and accuracy for matching of whole faces compared to matching of external or internal features. Although the results of Study 2 and Study 1 seem to be partially contradicting it is important to notice some differences in the methodological approaches chosen for the two studies.

In Study 2, and later in Study 3, the instruction of matching only whole face identity was no longer employed. Instead, subjects had to match faces only with respect to identity in external or internal features. Second, we have collapsed conditions ID and DF into condition IDE, and conditions SI and SE into condition SISE. Collapsing this way amounts to obtaining performance measures for whole-based decisions about face identity (ID and DF), and for part-based decisions about face identity (SI and SE), which motivates our data aggregation. Moreover, we have realised that whole-face matching strategies may be used whenever they are not precluded by the task, i.e. in IDE the subjects need not necessarily pay attention to only internal or external features, but may also exploit additional congruent context information available in IDE, which allows to resort to true whole based comparisons. This makes Task 1, where only whole faces had to be matched for identity, unnecessary. This is so because we can conclude possible whole-face processing strategies, enabled by the additional use of configural information, from a performance advantage in IDE compared to SISE. Matching of SISE faces is, however, expected to depend on the matching criteria as defined by instruction, and is expected to rely more on the information provided by internal or external features, respectively.

There are also methodological reasons for collapsing the four original categories ID, SI, SE and DF as proposed above. Since error rates are much more modulated by our independent variables than reaction times we need a bias-free measure of proportion correct (pc). Now, IDE is based on correct *same* (Hit) + correct *different* (Correct Rejection) responses for whole set judgements, while SISE is based on correct *same* (Hit) + correct *different* (Correct Rejection) responses for part set judgements. Calculating proportion correct for the four categories separately means that one obtains proportions which are biased by the subjects' tendency to say *same* or *different*. Hence, interpreting proportion correct for each the four categories is possible only if a subject responds bias-free.

The first two studies did not provide sufficient evidence for the unambiguous conclusion that there is an advantage of configural modes of processing compared to featural ones. The observed data generally suggested an advantage for global processes involving configural information, with whole faces being matched more accurately than internal or external features. However, there was an advantage in matching of external features compared to internal features, which is particularly pronounced if the task requires matching from different orientations or viewpoints, stressing the higher robustness of external features against viewing contingency. The results point into the direction of the dual-hypothesis proposed by Bruce (1988). According to Bruce (1988) an initial configural processing may take place so that overall information about a face can be obtained quickly. This global analysis works as a guideline, directing successive featural processing of more detailed aspects of faces, and global and featural processes work in parallel in order to garner further information about a face.

Although there are reasons to assume different routes for internal and external facial features, Study 1 and 2 leave us with the question at which time scales configural and featural processes occur, and how internal and external features possibly interact in order to produce a facial percept. Since revealing the time courses of featural and configural information requires rigorous masking techniques a new methodological approach was devised, which enabled to study the temporal succession of processing global and local facial aspects.

In Study 3 (Chapter 5) we measured proportion correct as functions of exposure duration with a masking paradigm, and alternating ¾ views of face stimuli. Matching performance thresholds were determined from psychometric curves for internal and external facial features in congruent and incongruent conditions (IDE vs. SISE). A similar methodological approach was chosen in the study by Grill-Spector and Kanwisher (2005), where percent correct was measured as a function of exposure duration in accomplishing three tasks: detection, categorization and identification of objects (Grill-Spector & Kanwisher, 2005). Their results showed that for object categorization and detection the same processing time is required. As soon as a subject could detect an object, he/she already knew its category (e.g. flower, car, dog). However, comparable performance on the identification task (e.g. rose, porsche, shepherd dog) required substantially longer time than either detection or categorization, which was accomplished within 50 ms. Similar results were obtained if the study by Sugase et al. (1999), where the activity of single neurons in the temporal cortex of macaque monkeys were recorded. Global information, categorizing stimuli as monkey faces, human faces, or shapes, was conveyed in the earliest part of the responses. Fine information about identity or expression was conveyed later, beginning on average 51 ms after the onset of global information. Both studies therefore imply fast processing for global aspects of objects and faces, which could be used as a 'header' to prepare destination areas for receiving more detailed information.

The results provided by our Study 3 are evidence for very fast processing of external features. Accuracy levels of 75% correct are reached with external features already within first 85 ms. Most importantly, we were able to show a striking performance asymmetry for internal and external features within the first 120 ms, suggesting that processing of whole faces in the early stream predominantly relies on external features, whereas internal features are added in later stages of face matching. As the comparison across the two instructions shows, the advantage of external features is a true perceptual effect, and is not due to attention being directed to internal or external features as a function of the task demands. The observed differences in the time courses do not imply serial processing stages for internal and external features, with external features coming first and internal features coming later. Instead our results indicate that both kinds of information evolve in time in parallel, but internal features with some delay. So, according to our observation that external features are available quick, and, after some time, both internal and external features are available also with a synergy effect in congruent contexts, it seems to be more founded to postulate two systems of processing. The first is fed by external features, and acts on a brief time scale. The second is fed by internal features and any configural information that establishes with the presence of both types of features. Synergy in congruent contexts and inhibition in conflicting contexts may be taken to indicate that both systems interact at later moments in time when information from both types of features can be integrated. In agreement with the latter proposal it was suggested that the object-recognition system and the face recognition system are involved in face perception, and interact to solve the task. The object recognition system is assumed to respond to global external information at a faster rate than the face-recognition system, which is assumed to respond to local internal information (Moscovitch & Moscovitch, 2000). At the end, and similar to the results of Study 1 and 2, we did find the best matching performance for whole faces, where internal and external features are congruent with respect to the correct response category.. However, results obtained in Study 3 have shown that this advantage emerges at later stages of face processing (after 200 ms), and depends on a successful decoding of internal facial information within the facial context.

Summarized, empirical evidence gained in our three studies favours the configural approach, where configuration has to be understood as a percept of the whole face given by a specific arrangement of its components (Bartlett & Searcy, 1993; Diamond & Carey, 1986; Rhodes et al., 1993; Searcy & Bartlett, 1996). Hereby configuration is not reduced only on spatial relations between internal features (as it is sometimes in the literature), but refers to the spatial arrangement of *all* features, i.e. relations among external features and internal features are an important part of configural information, enabling perception of complete 'wholes'. According to the configural approach, global aspect of faces are expected to be processed fast and accurate, whereas processing of facial features occurs part-by-part, and is usually considered as time consuming (for a review see Rakover, 2002). Although the distinction between configural and holistical approaches seems to be hardly detectable, the main difference emerges in the weighting of featural information. Applied to our studies where internal and external features were combined in natural facial stimuli, the two approaches lead to different hypotheses.

Both approaches propose matching at a global level as the first stage of perceptual processing. The holistical view, however, assumes that in this global representation featural or configural information is not made explicit, but assumes a facial Gestalt, perceived as an unimpaired perceptual wholeness (Ellis, 1975, Tanaka & Farah, 1993, Fahra et al., 1995; Carey & Diamond, 1994; Endo et al., 1989, Hole, 1994; Young et al., 1987). According to the holistical hypothesis, matching of whole faces would always (i.e. at any processing stage) be advantageous compared to matching of facial features. Similar results were obtained in studies with chimeric faces, where upper and lower halves of two original, different faces perceptually fuse to produce a strong impression of a complete novel face. It becomes difficult to perceive either half of the chimeric faces in isolation (Young et al., 1987; Carey & Diamond, 1994; Endo et al., 1989). The holistical approach would therefore imply a fusion of internal and external facial features also for our stimuli. However, we observe that the IDE condition leads to identical performance for brief timings in both tasks, matching of internal and matching of external features, but to clearly different curves for longer exposure durations, where higher performance was reached when internal features were to be matched. Further, matching of internal features suffers stronger from inversion, and at later moments in time, which is unexpected if featural and configural information are represented together, at one moment in time. Both observations are clear evidence against a holistical view.

On the other hand the configural approach proposes matching of global configural properties at an early stage, and matching of isolated features requiring scrutiny at the later stages if there is enough time to accomplish the feature-by-feature analysis (Hole, 1994). With regard to our stimuli this would imply accurate face matching already on the basis of external features, since external features transmit the most important aspect of the global structure. In later stages a time consuming feature-by-feature analysis works on isolated features, with the eyes region attracting about 60% of the whole processing time (Henderson, et al., 2001). This corresponds fairly well to our observation of flat psychometric curves for matching of internal features in the SISE condition, which enforces comparison of sets of internal features ignoring the external feature context, and steep and quickly saturating psychometric curves for matching of external features regardless of the degree of congruency, IDE or SISE.

6.2 The face recognition and the object recognition system

The large amount of evidence that face perception is mediated by special cognitive and neural mechanisms differing from those for objects comes from fMRI studies of the fusiform face area (FFA), and behavioural studies of the face inversion effect. There is strong evidence that inverting faces impairs the integration of features into a gestalt, the so-called holistic face representation (Sergent, 1984; Tanaka & Farah, 1993; Young et al., 1987). Many studies have shown that the perception of the relative positions of facial features within a face stimulus is more affected by inversion than the perception of local modifications applied to the facial features (Barton et al., 2001; Rhodes et al., 1993; Freire, Lee, & Symons, 2000). This led to the conclusion that the extraction of configural information is seriously impaired by face inversion.

Studying the face inversion effect (FIE) is, therefore, a relatively well established methodological approach to assess distinct processing paths for featural and configural information. We also exploited the modulation of the FIE in all three studies in order to reveal modes of featural and configural processing. There is evidence that, whereas the face recognition system forms holistic representations of faces based on configurations primarily of internal features, the object recognition system integrates information about individual features and the global relations among them, and is predominantly involved in the processing of external features (Moscovitch & Moscovitch, 2000; Moscovitch et al, 1997). Moreover, upright faces are found to be processed predominantly by the face recognition system, whereas processing of inverted presented facial stimuli demands strong activation of the object-recognition system (Moscovitch et al, 1997; Haxby et al., 1999; Farah et al, 1995). By exploring selective impairment in matching performance for inverted external and internal features within an intact facial context we were able to test the postulated differences in the two processing paths.

In Study 1 (Chapter 3) reaction times and error rates were compared in three matching tasks, focussing either on internal or external features, or on whole faces. A strong inversion effect was obtained for matching of whole faces (Task1), but not for facial features (Task 2 and 3). The data may be interpreted in favour of a holistical approach, proposing an inversion effect for whole faces only, but not for isolated features (for a review see Valentine, 1988). The lack of an inversion effect for both internal and external facial features may also be accounted as evidence against the different processing systems for external and internal facial features (Moscovitch & Moscovitch, 2000). However, there are good reasons why such conclusions should be considered with care. Firstly, matching performance with internal features was much worse than with external features, in both upright and inverted orientation, which is not expected if processing is mediated by a unique system. In addition, presenting target and test faces in frontal view, as done in Study 1, may also account for the observed pattern of data, since pictorial matching strategies were not precluded. This implies that it was not guaranteed that the subjects really extracted and compared facial features, but they could rely on judging the coincidence of image regions (Rakover & Teucher, 1997).

In the two following studies the experimental approach was toughly revised. Feature based matching was assured in Study 2 by presenting faces in different orientations, i.e. the first in upright and the second in inverted orientation, and in Study 3 by asking subjects to match mirrored faces in ¾ view. Both manipulations enforce subjects to match faces based on the observed features and their configural arrangement.

Indeed, in Study 2 (Chapter 4) and Study 3 (Chapter 5), where featural based matching was assured, external and internal facial features were found to suffer differently from inversion. The results obtained in Study 2 revealed a large effect of orientation for internal features, whereas for external features matching performance was nearly the same for upright-upright and upright-inverted matching pairs. Similarly, in Study 3 an inversion effect for matching of internal features was found at all exposure durations, remaining at a constant and high level for exposure durations beyond 200 ms. For matching of external features a FIE was identified for the shot exposure durations below 120 ms, but not for more relaxed timings. In both studies matching performance dropped significantly for inverted faces. Since a strong inversion effect for whole faces was obtained in Study 1 too, the results ratify that the configural or holistical facial percept was impaired by inversion, independent of the employed matching strategies in all three studies.

Not only the dependency of the FIE on feature type, but also the clear difference in the viewpoint-dependency (Study 2) and the timing of processing (Study 3) indicate different processing paths, and, behind it, separate processing systems for internal and external facial features. In Study 2 it was demonstrated that changes of viewpoint between target and test faces (from frontal view to ¾ view) caused large impairment of matching performance (increase in both error rates and reaction times) for internal features, whereas performance with external features was surprisingly unaffected by a variation of viewpoint.

In Study 3 we found evidence for fast processing of face stimuli when a subject's judgement could rely on just external feature information, and evidence for slower processing when judgements were based on purely internal features, and its inherent configural content. Fast matching with external features, independent of inversion or congruency, and the similarity to the time scales of object detection and categorization suggest that the object recognition system is involved at early face processing stages, leading to 75% correctness level in matching of external features within the first 85 ms, and to almost complete performance within the first 120 ms. Contrary, the face recognition system, which is activated by internal features and individual face differences (Bentin et al., 1996; Eimer, 2000; Jacques & Rossion, 2006; Jacques et al., 2007) acts on a slower time scale, and reaches the same levels of accuracy at the earliest after 160-300 ms (ibid.), which is the time scale found for matching of internal features.

The absence of reasonable performance with internal features within the first 120 ms, the well developed performance within this interval for external features, the high performance levels achieved with internal features after 200 ms, and the strong FIE for internal features for these longer exposure durations, accompanied by the absence of the FIE for external features, led to the conclusion that the object recognition system is activated by face outline and global shape within the first 120 ms, followed by an activation of the face recognition system, gradually starting after the first 120 ms, but being able to provide fine details and configural information when the first 200 ms have passed. The conclusion that there is a critical time window of 120 ms to 200 ms in which configural information is gradually supplied by the face recognition system is supported by current neuropsychological and neuroimaging findings (Farah et al., 1995; Haxby et al., 1999; Grill-Spector et al., 2000; Jacques & Rossion, 2006; Jacques et al., 2007), and is fully compatible with the set of results on the face inversion effect obtained so far.

7. Concluding Remarks and Suggestions for Future Research

Viewed together, the distinction between internal and external features has turned out to be quite fruitful for revealing featural and configural modes of processing in face perception. Combining this distinction with a proper experimental task that guarantees selective monitoring of both types of features in natural face stimuli it was possible to identify different temporal intervals for global featural information on the one hand, and detailed featural and configural information on the other. Since we preserved the natural face context there is a good degree of ecological validity for our basic findings, and the main result of a definite temporal order in which the two kinds of information become available in the stream of processing.

The validity of the present results could be enhanced with proper follow-up studies, aiming directly at the use of configural information. This could be done by taking a *developmental perspective*, and by further exploring the role of *facial contextual information* for processing of face parts. Further, combining timing measurements with fMRI recordings from object sensitive and face sensitive areas could also be a promising way of enhancing the validity of our findings.

a) Developmental perspective. If it can be shown that children and adults differ in the time scales of internal and external feature processing, and that the differential effects found for both types of features with adults are not found with children (or at timings very different from adults), this would be strong support that the relative weight in the use of featural and configural information in face recognition is subject to accumulating expertise. Existing studies give first hints that such an approach aiming at proving discriminant validity with extreme groups could be successful. Face processing changes during the childhood, and is supposed to reach an adult level at the age of 9 (Campbell & Tuck, 1995;

Mondloch et al., 2002). Hence, it is expected that the proportion of featural and configural face processing is different in children and adults, and also that facial feature processing runs on different times scales. Both aspects can be revealed with the methodology used in Study 3 to measure psychometric curves with a masking paradigm. In contrast, no differences are expected in two different adult age groups, since adults are supposed to be at the same level of expertise. Comparison of all three groups would be a critical test of discriminant validity for the definite temporal order of featural and configural information, as found here.

b) Facial contextual information. A major finding of Study 3 was that contextual information had a much stronger modulating effect on judging of internal features than judging of external features. For judging of internal features, congruent contexts were helpful, while incongruent contexts let task performance deteriorate. This indicates that judgements about features which contain configural information are subject to context information, which establishes new sources of configural information that may be used in the task.

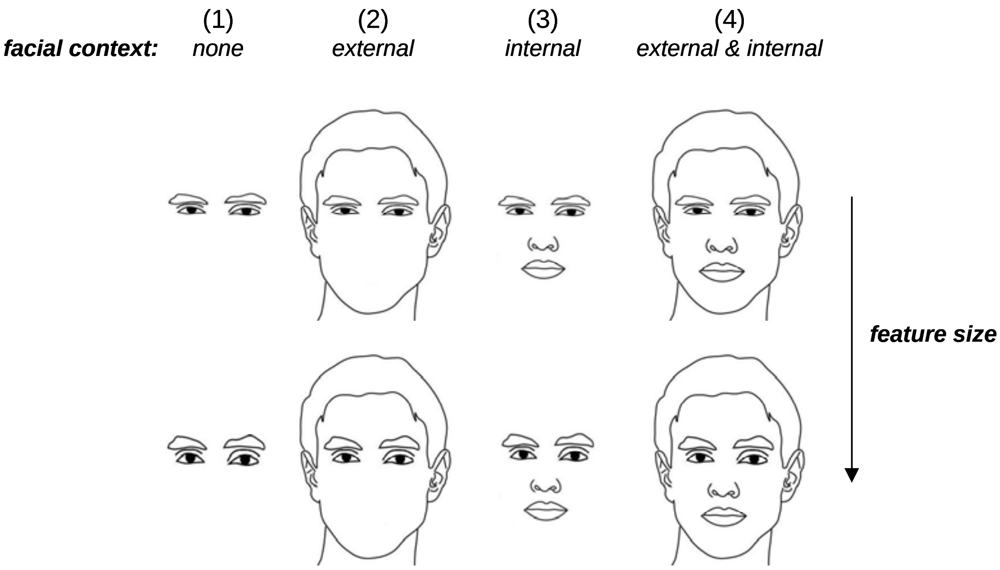

Figure 7.1: Stimulus example for a follow-up study on the effects of different degrees of contextual information on facial feature size judgements. The size of eyes and eyebrows is scaled in the target stimuli (lower row) relative to the reference (upper row). Four facial contexts are used: no context, external features, internal features and both (= whole face context).

A follow-up study dedicated to the use of configural information could build on this basic finding of Study 3, asking how a judgement subject to configural information is modulated by different degrees of facial context information. A possible task is shown in Figure 7.1. In the experiment the just noticeable size difference for pairs of eyes and eyebrows is measured in a 2AFC task. This task is performed with: (1) isolated features, (2) external feature context, (3) internal feature context, and (4) external and internal feature context (see Figure 7.1). Since size judgements can be expected to exploit changes of distances to other features which naturally occur as the feature size is altered, the experiment reveals which types of feature context are most effective, and how large the degree of summation becomes when both kinds of feature contexts are present. Since, again, timing functions can be measured, contextual modulation of internal and external features can be judged not only with respect to accuracy, but also with respect to time course.

c) fMRI-validation. The timing of featural and configural information processing found in the present study can be validated with proper fMRI studies. Recording from the object sensitive LOC and the face sensitive fusiform face area (FFA), a differential scheme of activation at different moments in time should emerge. Independent of instruction there should be activation in the LOC in the first 120 ms of processing. Dependent on instruction, activation should enter in the FFA to different degrees: FFA activation should become stronger, but gradually fade in the LOC for monitoring internal features. For monitoring external features a strong LOC activation should reside at longer exposure durations, and only a minor FFA activation should emerge. Indeed, executing our basic experimental paradigm while simultaneously recording from both areas would be an ideal means for validation of the temporal precedence of global feature information, and the critical time window of 120 ms to 200 ms for establishing detailed featural and configural information in the course of face processing. This is left to future projects.

8. References

1. Aguirre, G. K., Zarahn, E., & D`Esposito., M. (1998). An area within human ventral cortex sensitive to `building`stimuli: evidence and implications. *Neuron, 21*, 373-383.
2. Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Acadademy of Science USA, 90*, 5718–5722.
3. Ahissar, M., & Hochstein, S. (1996) Learning pop-out detection: Specifities to stimulus characteristics. *Vision Research, 36*, 3487–3500.
4. Ahissar, M., & Hochstein, S. (1997) Task difficulty and the specifity of perceptual learning. *Nature, 378*, 401–406.
5. Bachmann, T. (1991). Identification of spatially quantised tachistostopic images of faces: How many pixels does it take to carry identity?, *European Journal of Cognitive Psychology, 3*, 87-103.
6. Baddeley, A., & Woodhead, M. (1983). Improving face recognition ability. In S. Llooyd-Bostock & B. Clifford, *Evaluating witness evidence* (pp.125-136). Chichester: Wiley.
7. Bartlett, J. C., & Searcy, J. (1993). Inversion and configurational of faces. *Cognitive Psychology, 25*, 281-316.
8. Bartlett, J. C., Searcy, J. H, & Abdi, H. (2003). What are the routes to face recognition? In M.A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp.21-52). Oxford: Oxford University Press.
9. Bartlett, J. C., Hurry, S., & Thorley, W. (1984). Typicality and familiarity of faces. *Memory and Cognition, 12,* 219-228.
10. Barton, J. J., Keenan, P. J., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effect of inversion and viewing duration. *British Journal of Psychology, 92*, 527-549.
11. Benson, P. J., & Perrett, D. I. (1991). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology, 3,* 105-135.
12. Benson, P. J., & Perrett, D. I. (1994). Visual processing of facial distinctiveness. *Perception, 23,* 75-93.
13. Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8,* 551–565.
14. Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115-147.
15. Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research, 39,* 2885-2899.
16. Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint

invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162-1182.
17. Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London, 352,* 1203-1219.
18. Blanz, V., Tarr, M. J., & Bülthof, H. H. (1999). What object attributes determine canonical views? *Perception, 28,* 575-599.
19. Bodamer, J. (1947). Die Prosopagnosie. *Archiv für Psychiatrie und Nervenkrankheiten, 179,* 6-53.
20. Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *5*, 527-536.
21. Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception and Psychophysics, 65,* 1078-1093.
22. Bradshaw, J. L., & Wallace, G. (1971). Models for the processing and identification of faces, *Perception and Psychophysics, 9*, 443-448.
23. Breitmeyer, B. G. (1984). *Visual Masking: An integrative Approach.* New York: Oxford University Press.
24. Brennan, S. E. (1985). The caricature generator. *Leonardo, 18,* 170-179.
25. Bruce, V. (1988). *Recognizing faces*. London: Erlbaum.
26. Bruce, V., Burton, A. M., & Dench, N. (1994). What`s distinctive about a distinctive face? *Quarterly Journal of Experimental Psychology, 47A,* 119-141.
27. Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B, Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*, 339-360.
28. Bruce, V., Valentine, T, & Baddeley, A. D. (1987). The basis of the ¾ view advantage in face recognition. *Applied Cognitive Psychology, 1,* 109-120.
29. Bruck, M., Cavanagh, P., & Ceci, S. J. (1991). Fortysomething: Recognizing faces at one's 25th reunion. *Memory and Cognition, 19*, 221–228.
30. Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprints experts. *Vision Research, 45,* 431-448.
31. Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view theory of object recognition. *Proceedings of the National Academy of Science, USA, 89*, 60-64.
32. Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 3,* 247-260.
33. Campbell R., & Tuck, M. (1995). Recognition of parts of famous-face photographs by children: an experimental note. *Perception, 24,* 451 – 456.
34. Carbon, C. C., & Leder, H. (2005). When feature information comes first! Early processing of inverted faces. *Perception, 34,* 1117-1134.

35. Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195,* 312-314.
36. Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1,* 253-274.
37. Cohen, M. E., & Carr, W. J. (1975). Facial recognition and the von Restorff effect. *Bulletin of the Psychonomic Society, 6,* 383-384.
38. Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processing in the recognition of faces of different familiarity. *Perception, 29*, 893-909.
39. Coltheart, M. (1977). Contemporary models of the cognitive processes. I. In V. Hamilton & M. D. Vernon (Eds.), *Information input and storage. The development of cognitive processes* (pp.11-41). London: Academic Press.
40. Coltheart, M. (1983). Iconic memory. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 302*, 283-294.
41. Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics, 58,* 602-612.
42. Davies, G., Ellis, H., & Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology, 63,* 180-187.
43. Davies, G., & Milne, A. (1982). Recognizing faces in and out of context. *Current Psychology, 4*, 235-246.
44. De Haan, E. H. F., & Hay, D. C. (1986). The matching of famous and unknown faces, given either the internal or the external features: A study of patients with unilateral brain lesions. In H. E. Ellis, M. A. Jeeves, F. Newcombe, & A. W. Young (Eds.), *Aspects of face processing* (pp.302-309). Dordrecht The Netherlands: Martinus Niijhoff.
45. Diamond, R., & Carey, S. (1986). Whay faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107-117.
46. Edelman, S. (1995). Class similarity and viewpoint invariance in the recognition of 3-D objects. *Biological Cybernetics, 72,* 207-220.
47. Edelman, S., & Bülthoff, H.H. (1992). Orientation dependence in the recognition of familar and novel views of three-dimensional objects. *Vision Research, 32,* 2385-2400.
48. Eger, E., Jedynak, A., Iwaki, T., & Skrandies, W. (2003). Rapid extraction of emotional expression: evidence form evoked potential fields during brief presentation of face stimuli. *Neuropsychologia, 41,* 808-817.
49. Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport, 11,* 2319-2324.
50. Eley, M. G. (1982). Identifying rotated letter-like symbols. *Memory and Cognition, 10,* 25-32.
51. Ellis, H. D. (1975). Recognizing faces. *British Journal of Psychology, 66,* 409-426.

52. Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, 8, 431-439.
53. Endo, M., Masame, K., & Maruyama, K (1989). Interference from configuration of a schematic face onto the recognition of its constituent parts. *Tohoku Psychologica Folia, 48,* 97-106.
54. Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *NeuroReport*, 9, 303-308.
55. Fagan, J. (1979). The origins of facial pattern recognition. In M. H. Bornstein & W. Kessen, *Psychological developement from infancy: image to intention* (pp.83-113). Hillsdale, NJ: Lawrence Erlbaum Associates.
56. Farah, M. J. (1996). Is face recognition special? Evidence from neuropsychology. *Behavioral Brain Research, 76,* 181-189.
57. Farah, M. J., Rochlin, R, & Klein, K. L. (1994). Orientation invariance and geometric primitives in shape recognition. *Cognitive Science, 18,* 325-344.
58. Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628-634.
59. Farah, M. J., Wilson, K., D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35,* 2089-2093.
60. Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception, 12,* 195-201.
61. Flavell, J. H., & Draguns, J. (1957). A microgenetic approach to perception and thought. *Psychological Bulletin, 54,* 197-217.
62. Fraser, I. H., Craig, G. L., & Parker, D. M. (1990). Reaction time measures of features saliency in schematic faces. *Perception, 19,* 661-673.
63. Freire, A., Lee, K., & Symons, L. A. (2000). The face inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29,* 159-170.
64. Frowd, C., Bruce, V., McIntyre, A., Hancock, P. J. (2007). The relative importance of external and internal features of facial composites. *British Journal of Psychology, 98,* 61-77.
65. Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science, 212,* 952-955.
66. Galton, F. (1879). Composite portraits, made by combining those of many different persons into a single, resultant figure. *Journal of Anthropological Insitute, 8,* 132-144.
67. Garner, W. R. (1978). Aspects of a stimulus: Features, dimensions, and configurations. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp.99-133). Hillsdale, NJ: Erlbaum.

68. Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training “greeble” experts: a framework for studying expert object recognition processes. *Vision Research, 38,* 2401-2428.
69. George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research, 4,* 65–76.
70. Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception, 34,* 77-86.
71. Goffaux, V., & Rossion, B. (2006). Faces are “spatial”- holistic perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1023-1039.
72. Going, M., & Read, J. D. (1974). The effect of uniqueness, sex of participant and sex of photograph on the facial recognition. *Perceptual and Motor Skills, 39,* 109-110.
73. Goldstein, A. G. (1965). Recognition of inverted photographs of faces in children and adults. *Psychonomic Science, 3,* 447-448.
74. Green, D. M., Swets, J. A. (1988). *Signal detection theory and psychophysics.* Wiley, Los Altos.
75. Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16,* 152-160.
76. Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object–selective activation correlate with recognition performance in humans. *Nature Neuroscience,* 3, 837-843.
77. Gross, C. G., Roche-Miranda, G. E., & Bender, D. B. (1972). Visual properties of neurons in the inferotemporal cortex of macaque. *Journal of Neurophysiology, 35,* 96-111.
78. Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Science, 4,* 330-337.
79. Harmon, L. D. (1973). The recognition of faces. *Scientific American, 227,* 71-82.
80. Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science, 180,* 1194-1197.
81. Harries, M. H., & Perrett, D. I. (1991). Visual processing of faces in temporal cortex: Physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience, 3,* 9-24.
82. Harries, M. H., Perret, D. I., & Lavender, A. (1991). Preferential inspection of views of 3-D model heads. *Perception, 20,* 669-680.
83. Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expresion and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research, 32,* 203-218.

84. Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233.
85. Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron, 22,* 189-199.
86. Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (Ed.) *Normality and Pathology in Cognitive Function* (pp.173-202). London: Academic Press.
87. Hayward, W. G. (2003). After the viewpoint debate: where next in object recognition? *Trends in Cognitive Sciences, 7,* 425-427.
88. Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1511-1521.
89. Hayward, W. G., & Williams, P. (2000). Viewpoint dependence and object discriminability. *Psychological Science, 11,* 7-12.
90. Henderson, J. M., Falk, R. J., Minut, S., Dyer, F. C., & Mahadevan, S. (2001). Gaze control for face learning and recognition by humans and machines. In T. Shipley & P. Kellman (Eds), *From fragments to objects: Segmentation processes in vision*. (pp.463-482). Amsterdam: Elsevier.
91. Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception & Performance, 22,* 986-1004.
92. Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62,* 201-222.
93. Hines, D., Jordan-Brown, L., & Juzwin, K. R. (1987). Hemispheric visual processing in face recognition. *Brain and Cognition, 6,* 91–100.
94. Hochberg, J., & Galper, R. E. (1967). Recognition of faces: 1. An exploratory study. *Psychonomic Science, 9,* 619-620.
95. Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23,* 65-74.
96. Hole, G. J., George, P. A., & Dunsmore, V. (1999). Evidence for holistic processing of faces viewed as photographic negatives. *Perception, 28,* 341-359.
97. Hoeger, R. (1997). Speed of processing and stimulus complexity in low-frequency and high-frequency channels. *Perception, 26,* 1039-1045.
98. Hubner, R. (1997). The effect of spatial frequency on global precedence and hemispheric differences. *Perception and Psychophysics 59,* 187-201.
99. Humphreys, G. W, & Riddoch, M. J. (1984). Routes to object constancy: Implications from neurological impairments of object constancy. *Quarterly Journal of Experimental Psychology A, 37,* 469-495.
100. Intraub, H. (1999). Understanding and remembering briefly glimpsed pictures: Implications for visual scanning and memory. In Coltheart, V.

(Ed.), *Fleeting Memories: Cognition of Brief Visual Stimuli* (pp.47-70), Cambridge, Massachusetts: MIT Press.

101. Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceeding of the National Academy of Sciences of the United States of America, 96,* 9379–9384.
102. Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7,* 1-9.
103. Jacques, C., & Rossion, B. (2006). The speed of individual face categorization., *Psychological Science, 17,* 485–492.
104. Jarudi, I. N., & Sinha, P. (2003). Relative contributions of internal and external features to face recognition. *Massachusetts Institute of technology, Tech.Rep.*
105. Jeffreys, D. A. (1989). A face responsive potential recorded from the human scalp. *Experimental Brain Research,* 78, 193-202.
106. Johnston, R. A., & Ellis, H. D. (1995a). Age effects in the processing of typical and distinctive faces. *Quarterly Journal of Experimental Psychology, 48A,* 447-465.
107. Johnston, R. A., & Ellis, H. D. (1995b). The development of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of faces recognition – explorations in face space* (pp. 1-23). London: Routledge.
108. Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision,* 3, 499-512.
109. Johnson, R. A., & Wichern, D. W. (2003). *Applied multivariate statistical analysis.* Prentice Hall, Upper Saddle River (New Jersey).
110. Kanwisher, N. G., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302-4311.
111. Kanwisher, N. G., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London, B, 361,* 2109-2128.
112. Kemp, R., McManus, C., & Pigott, T. (1990). Sensitivity to the displacement of facial features in negative and inverted images. *Perception, 19,* 531-543.
113. Keysers, C., Xiao, D. K., Foldiak, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience, 13,* 90-101.
114. Keysers, C., Xiao, D. K., Foldiak, P., & Perrett, D. I. (2005). Out of sight but not out of mind: The neurophysiology of iconic memory in the superior temporal sulcus. *Cognitive Neuropsychology, 22,* 316-332.
115. Kosslyn, S. M. (1994). *Image and brain: the resolution of the imagery debate.* Cambridge, MA: MIT Press.

116. Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66,* 651-654.
117. Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proceedings of the National Academy of Science, USA, 92,* 5587-5591.
118. Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1,* 115-121.
119. Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science, 5,* 214-220.
120. Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica, 102,* 221-245.
121. Leder, H., & Burce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology, 51A,* 449-473.
122. Leder, H., & Burce, V. (2000). When inverted faces are recognized: the role of configural information in face recognition. *Quarterly Journal of Experimental Psychology, 53A,* 513-536.
123. Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30,* 73–83.
124. Lee, Y., Matsumiya, K., & Wilson, H. R. (2006). Size-invariant but viewpoint-dependent representation of faces. *Vision Research, 46,* 1901-1910.
125. Lehky, S. R. (2000). Fine discrimination of faces can be performed rapidly. *Journal of Cognitive Neuroscience, 12,* 848-855.
126. Lewis, M. B., Johnston, R. A. (1997). The Thacher illusion as a test of configural disruption. *Perception, 26,* 225-227.
127. Light, L. L., Kayra-Stuart, F, & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 212-228.
128. Lindsay, D. S., Jack, P. C., Christian, M. A. (1991). Other-race face perception. *Journal of Applied Psychology,* 76, 587-589.
129. Liu, C. H., & Chaudhuri, A. (1998). Are there qualitative differences in face recognition between positive and negative? *Perception, 27,* 1107-1122.
130. Liu, C. H., & Chaudhuri, A. (2002) Reassessing the ¾ view effect in face recognition. *Cognition, 83,* 31-48.
131. Liu, C. H., Collin, C. A., Burton, A. M., & Chauduri, A. (1999). Lighting directions affects recognition of untextured faces in photographic positive and negative. *Vision Research, 39,* 4003-4009.
132. Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience, 5,* 910-916.

133. Logie, R. H., Baddeley, A. D., & Woodhead, M. M. (1987). Face recognition, pose and ecological validity. *Applied Cognitive Psychology, 1,* 53-69.
134. Love, B. C., Rouder, J. N., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology, 38,* 291-316.
135. Lowe, D. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31,* 355-395.
136. Macho, S., & Leder, H. (1998). Your eyes only? A test of interactive influence in processing of facial features. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1486-1500.
137. Malcolm, L. C., Leung, C., & Barton, J. J. (2005). Regional variation in the inversion effect for faces: Differential effects of feature shape, feature configuration, and external contour. *Perception, 34,* 1221-1231.
138. Marr, D. (1982). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information.* Freeman, New York.
139. Marr, D., & Nishihara, H. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceeding of the Royal Society of London B, 200,* 269-294.
140. Martens, U., Schweinberger, S. R., Kiefer, M., & Burton, A. M. (2006). Masked and unmasked electrophysiological repetition effects of famous faces. *Brain Research, 1109,* 146-157.
141. Mauro, R., & Kubovy, M. (1992). Caricature and face recognition. *Memory and Cognition, 20,* 433-440.
142. Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery and Psychiatry, 37,* 489-501.
143. McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9,* 605-610.
144. McKelvie, S. (1976). The role of eyes and mouth in the memory of a face. *American Journal of Psychology, 2,* 311-323.
145. Miyashita, Y. (1988). Neuronal correlates of visual associative long-term memory in the primate temporal cortex. *Nature, 335,* 817-820.
146. Miyashita, Y., & Chang, H. S. (1988). Neural correlates of pictorial short-term memory in the primate temporal cortex. *Nature, 331,* 68-70.
147. Mondloch, C. J., Dobson, K. S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds cannot tell the difference between Steve Martin and Paul Newman: factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology, 89,* 159-181.
148. Mondloch, C. J., Grand, L. R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31,* 553 – 556.

149. Moore, C., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *Journal of Neuroscience, 26,* 11187-11196.
150. Moscovitch, M., & Moscovitch, D. A. (2000). Super face-inversion effects for isolated internal or external features, and for fractured faces. *Cognitive Neuropsychology, 17,* 201–219.
151. Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What Is Special about Face Recognition?: Nineteen Experiments on a Person with Visual Object Agnosia and Dyslexia but Normal Face Recognition. *The Journal of Cognitive Neuroscience, 9,* 555 -604.
152. Nachson, I., Moscovitch, M., & Umilta, C. (1995). The contribution of external and internal features to the matching of unfamiliar faces. *Psychological Research 58,* 31-37.
153. Nachson, I., & Shechory, M. (2002). Effect of inversion on the recognition of external and internal facial features. *Acta Psychologica, 109,* 227-238.
154. Nagayama, R., Yoshida, H., & Toshima, T. (1995). Interrelationship between the facial expression and familiarity: Analysis using spatial filtering and inverted presentation. *Japanese Journal of Psychology, 66,* 327-335.
155. Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.
156. Navon, D. (1981). Do attention and decision follow perception? Comment on Miller. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1175-1182.
157. Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39,* 3824-3833.
158. Newell, F. N. (1998). Stimulus context and view dependence in object recognition. *Perception, 27,* 47-68.
159. Newell, F. N., Chiroro, P., & Valentine, T. (1999). Recognizing unfamiliar faces. The effect of distinctiveness and view. *Quarterly Journal of Experimental Psychology, Section A, 52,* 509-534.
160. Op de Beek, H., Baker, C., DiCarlo, J, & Kanwisher, N. (2006). Discrimination Training Alters Object Representations in Human Extrastriate Cortex. *Journal of Neuroscience, 26,* 13025-13036.
161. O`Toole, A. J., Edelman, S., & Bülthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38,* 2351-2363.
162. Palmer, S. (1983). The psychology of perceptual organization: A transformational approach. In J. Beck, B. Hope & A. Rosenfeld (Eds.). *Human and machine vision* (pp. 269-340). New York: Academic Press.
163. Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long and A. Baddeley (Eds.), *Attention and Performances IX* (pp. 135-151). Lawrence Erlbaum, Hillsdale, NJ.

164. Patterson, K. E., & Baddeley, A. D. (1977). When faces recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 406-417.
165. Perret, D. I., Mistlin, A. J., & Chitty, A. J. (1989). Visual neurones responsive to faces. *Trends in Neuroscience, 10,* 358-364.
166. Perret, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research, 86,* 159-173.
167. Perret, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47,* 329-342.
168. Phillips, R. J. (1979). Some exploratory experiments on memory for photographs of faces. *Acta Psychologica, 43,* 39–56.
169. Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature, 343,* 263-266.
170. Poggio, T., & Vetter, T. (1992). Recognition and structure from 2d model view: Observations on prototypes, object classes and symmetries. *A.I. Memo No. 1347, Artifical Intelligence Laboratory; Massachusetts Institute of Technology.*
171. Press,W. H., Teukolsky, S. A., Flamery, B. P., Vetterling,W. T. (1996). *Numerical Recipes in Pascal.* Oxford University Press.
172. Puce, A., Allison, T., Asgari, M., Gore, J. C., McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16,* 5205-5215.
173. Rakover, S. S. (2002). Featural vs. configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology, 93,* 1-30.
174. Rakover, S. S., & Teucher, B. (1997). Face inversion effects: Parts and whole relationship. *Perception and Psychophysics 59,* 752 – 761.
175. Rhodes, G. (1988). Looking at faces: first-order and second-order features as determinants of facial appearance. *Perception, 17*, 43-63.
176. Rhodes, G., Brake, S., & Atkinson, A. (1993). What`s lost in inverted faces? *Cognition, 47,* 25-57.
177. Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and rating of caricatures: Implications for mental representations of faces. *Cognitive Psychology, 19,* 473-497.
178. Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the Fusiform Face Area Specialized for Faces, Individuation, or Expert Individuation? *Journal of Cognitive Neuroscience, 16,* 189-203.
179. Rhodes, G., Carey, S., Byatt, B., & Proffitt, F. (1998). Coding spatial variations in faces and simple shapes: A test of two models. *Vision Research, 38,* 2307-2321.

180. Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogenous stimuli: Toward a model of configural coding. *Perception*, *19*, 773-794.
181. Rhodes, G., & Tremewan, T. (1996). Understanding face recognition: caricature effects, inversion and the homogeneity problem. *Visual Cognition, 1,* 275-311.
182. Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience, 3,* 1199-1204.
183. Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception, 1988,* 17, 475-481.
184. Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contribution to theories of part/whole organization. *Cognitive Psychology, 23,* 299-330.
185. Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioral tasks. *Cognition, 103,* 34-79.
186. Rock, I., & Di Vita, J. (1987). A case of viewer-centered object representation. *Cognitive Psychology, 19,* 280-293.
187. Rolls, E. T., & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society of London B: Biological Sciences*, *257,* 9-15.
188. Rossion, B, & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision, 8,* 1-13.
189. Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21,* 329-330.
190. Schiller, P. H. (1968). Single unit analysis of backward visual masking and metacontrast in the cat lateral geniculate nucleus. *Visual Research, 8,* 855-866.
191. Schwaninger, A., Carbon, C. C., & Leder, H. (2003). Expert face processing: Specialization and constraints. In G. Schwarzer & H. Leder (Eds.), *Development of face processing* (pp.81-97*)*, Göttingen: Hogrefe.
192. Schwaninger, A., & Mast, F. W. (2005) The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism. *Japanese Psychological Research, 47,* 216-222.
193. Schwarzer, G, & Zauner, N. (2003). Face processing in 8-month-old infants: evidence for configural and analytical processing. *Vision Research 43*, 2783-2793.
194. Schyns, P. G., & Bülthff, H. H. (1994). Viewpoint dependence and face recognition. In A. Ram & K. Eiselt (Eds.), *Proceedings of the XVI meeting of the Cognitive Science Society* (pp. 789-793). Hillsdale, NJ: Erlbaum.
195. Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial relational information in faces. *Journal of*

Experimental Psychology: Human Perception and Performance, 22, 904-915.

196. Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., Landis, T., & Schomer, D. L. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport, 8*, 2749-2754.
197. Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75,* 221-242.
198. Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain, 115,* 15-36.
199. Shepherd, R. N., & Cooper, L. A. (1982). *Mental images and their transformations.* MIT Press, Cambrige, MA.
200. Shepherd, J., Davies, G, & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis & J. Shepherd (Eds.), *Perceiving and Remembering Faces* (pp. 105–131). London: Academic Press.
201. Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.
202. Smith, E. E., & Nielsen, G. D. (1970). Representation and retrival processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology, 85,* 397-405.
203. Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature, 369,* 395-397.
204. Sternberg, S. (1969). The discovery of processing stages: Extension of Donders`method. *Acta Psychologica, 30,* 276-315.
205. Sternberg, S. (1998a). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Methods, models ad conceptual results* (pp. 703-863). Cambridge, MA: MIT Press.
206. Sternberg S. (1998b). Inferring mental operations from reaction-time data: How we compare objects. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Methods, models ad conceptual results* (pp. 365-454). Cambridge, MA: MIT Press.
207. Stevenage, S. V. (1995). Can caricature really produce distinctiveness effect? *British Journal of Psychology, 86,* 127-146.
208. Sugase, Y., Yamane, S., Ueno, S, & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400,* 869-873.
209. Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 45,* 34-79.
210. Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and cognition, 25*, 583-592.

211. Tarr, M. J. (1995). Rotating object to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychological Bulletin and Review, 2,* 55-82.
212. Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233-282.
213. Tarr, M. J., & Pinker, S. (1990). When does human object recognition use a viwer-centered reference frame? *Psychological Science, 1,* 253-256.
214. Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9,* 483-484.
215. Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.
216. Toyama, J. S. (1975). The effect of orientation on the recognition of faces: A reply to Yin. *Unpublished PhD thesis, University of Waterloo, Canada.*
217. Treisman, A., & Gelade G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.
218. Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. *Vision Research, 36,* 1761-1771.
219. Tversky, A., & Krantz, D. H. (1969). Similarity of schematic faces: A test of interdimensional additivity. *Perception and Psychophysics, 5,* 124-128.
220. Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32,* 193-254.
221. Ullman, S., & Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13,* 992-1005.
222. Valentine, T. (1988). Inverted faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471-491.
223. Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology, Section A, 43*(2), 161 - 204 .
224. Valentin, D., Abdi, H., & Edelman, B. (1997). What represents a face? A computational approach for the integration of physiological and psychological data. *Perception, 26,* 1271-1288.
225. Valentine, T., & Bruce, V. (1986a). The effects of distinctiveness in recognising and classifying faces. *Perception, 15,* 525-535.
226. Valentine, T., & Bruce, V. (1986b). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica, 61,* 259-273.
227. Valentine, T., & Bruce, V. (1986c). Recognizing familiar faces: The role of distinctiveness and familiarity. *Canadian Journal of Psychology, 40,* 300-305.
228. Van Lier, R., & Wagemans, J. (1999). From images to objects: global and local completions of self-occluded parts. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1721-1741.

229. Vanrie, J., Béatse, E., Wagemans, J., Sunaert, S., & Van Hecke, P. (2002). Mental rotation versus invariant features in object perception from different viewpoints: an fMRI study. *Neuropsychologia, 40,* 917–930.
230. Vanrie, J., Willems, B., & Wagemans, J. (2001). Multiple routes to object matching from different viewpoints: mental rotation versus invariant features. *Perception, 30,* 1047-1056.
231. Vecera, S. P. (1998). Visual object representation: An introduction. *Psychobiology*, *26*, 281-308.
232. Veres-Injac, B., & Scwaninger, A. (in press). The time course of processing external and internal features of unfamiliar faces. *Psychological Research*.
233. Veres-Injac, B., & Scwaninger, A. (submitted). Does matching of internal and external facial features depend on viewpoint? *Acta Psychologica*.
234. Verfaillie, K. (1992). Variant points of view on viewpoint invariance. *Canadian Journal of Psychology, 46*, 215-235.
235. Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception and Psychophysics, 57*, 925-61.
236. Walker-Smith, G. J. (1978). The effect of delay and exposure duration in a face recognition task. *Perception and Psychophysics, 24*, 63-70.
237. Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6*, 313-326.
238. Wallis, G. M., & Bülthoff, H. H. (1997). Temporal correlations in presentation order during learning affects human object recognition. *Perception, 26*. ECVP Abstract Supplement.
239. Wallis, G. M., & Rolls, E. T. (1997). A model of invariant object recognition in the visual system. *Progress in Neurobiology, 51*, 167-194.
240. Wenger, M. J., & Townsend, J. T. (2000). Spatial frequencies in short-term memory for faces: A test of three frequency-dependent hypotheses. Memory and *Cognition, 28,* 125-142.
241. Williams C. C., & Henderson, J. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & Cognition, 35,* 1977–1985.
242. Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 181-190.
243. Woodhead, M., Baddeley, A. D., & Simmons, D. C. V. (1979). On training people to recognize faces. *Ergonomics, 22,* 333-343.
244. Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey. *Experimental Brain Research*, *73,* 209-214.
245. Yarmey, A. D. (1971). Recognition memory for familiar "public" faces: Effect of orientation and delay. *Psychonomic Science, 24,* 286-288.
246. Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141-145.

247. Yin, R. K. (1970). *Face recognition: A special process?* Unpublished PHD thesis, MIT.
248. Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. V. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*, 737-746.
249. Young, A.W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747-759.
250. Yue, X., Tjan, B., & Biederman, I. (2006). What makes faces special? *Vision Research, 46,* 3802-3811.

Curriculum Vitae

Personal Information

Veres-Injac, Bozana
Date of Birth: 10 August 1977, in Mrkonjic-Grad, Bosnia and Herzegovina

Education

09/1992 - 06/1996	High school in Novi Sad, Serbia (Typus B)
10/1996 - 05/2001	Study of Psychology at the University Novi Sad, Serbia, Department of Psychology. Diploma thesis: *The effect of subordinate and superordinate class in categorization tasks.*
10/2002 – 12/2005	Master study at the University Belgrad, Serbia, Department of Psychology, General Psychology. Master thesis: *The effect of stimulus exposure in semantic verification tasks.*
01/2006 – 09/2008	Doctoral study at the University Zurich, Department of Psychology, General Psychology (Cognition). Leader: Prof. Dr. W. Marx.